

UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

ANÁLISE ECONOMETRICA DE EFICIÊNCIA TÉCNICA
USANDO REGRESSÃO CANÔNICA NA ESTIMAÇÃO
DA FRONTEIRA ESTOCÁSTICA DE PRODUÇÃO

Tese submetida à Universidade Federal de Santa Catarina para a obtenção
do título de Doutor em Engenharia de Produção

JOÃO SERAFIM TUSI DA SILVEIRA

FLORIANÓPOLIS
2000

JOÃO SERAFIM TUSI DA SILVEIRA

**ANÁLISE ECONOMETRICA DE EFICIÊNCIA TÉCNICA
USANDO REGRESSÃO CANÔNICA NA ESTIMAÇÃO
DA FRONTEIRA ESTOCÁSTICA DE PRODUÇÃO**

Esta tese foi julgada adequada para a obtenção do título de DOUTOR EM
ENGENHARIA DE PRODUÇÃO e aprovada em sua forma final pelo
Programa de Pós-Graduação em Engenharia de Produção

Prof. Ricardo Miranda Barcia, Ph.D.
Coordenador do Curso

BANCA EXAMINADORA:

Prof. Edgar Augusto Lanzer, Ph.D.
Orientador

Prof. Helson Cavalcanti Braga, Dr.
Examinador Externo

Prof. João Rogério Sanson, Ph.D.

Prof. Pedro Alberto Barbetta, Dr.
Moderador

Prof. Ernani Hickmann, Ph.D.
Examinador Externo

Florianópolis, 14 de Fevereiro de 2000

À Maria do Rozário, minha mulher amada.

Aos meus queridos filhos Rafael, Diego e Letícia.

Aos meus pais Bibiano (in memorian) e Erondina.

Ao vô Bicaca e à vó Lota.

AGRADECIMENTOS

Na realização deste trabalho contei com o apoio de algumas pessoas e instituições a quem sou profundamente grato. Em especial,

à minha família pela compreensão e carinho com que aceitou os momentos subtraídos de nossa convivência;

ao professor Edgar A. Lanzer, orientador e amigo, pelo permanente incentivo e contribuições ao desenvolvimento da pesquisa;

ao professor e amigo Pedro Alberto Barbetta, pela atenção e competência com que ofereceu importantes sugestões sobre os métodos estatísticos não-paramétricos;

aos examinadores externos, cuja participação na banca elevou sobremaneira o evento e o Programa de Pós-Graduação em Engenharia de produção da UFSC;

ao professor William H. Greene do Departamento de Economia da Universidade de Nova York, pela acolhida e pelos preciosos conhecimentos transmitidos;

aos professores Tim Coelli (*University of New England, Australia*), Shawna Grosskopf (*Southern Illinois University, USA*), John Ruggiero (*University of Dayton, USA*) e Kwabena Gyimah-Brempong (*University of South Florida, USA*), pelas boas sugestões que enviaram por e-mail;

à UFSC que, através do Departamento de Ciências Econômicas, facultou dedicação integral ao empreendimento;

à Fundação de Estudos e Pesquisas Sócio-Econômicas (FEPES) da UFSC, pelo apoio emprestado para a realização de estudos no exterior e

ao Serviço de Controle Financeiro do Hospital Universitário da UFSC, pelo apoio na reprografia.

SUMÁRIO

LISTA DE TABELAS	VII
LISTA DE FIGURAS	VIII
LISTA DE APÊNDICES	IX
LISTA DE SIGLAS	X
LISTA DOS PRINCIPAIS SÍMBOLOS E ABREVIATURAS	XII
RESUMO	XVI
ABSTRACT	XVII
1 INTRODUÇÃO	1
1.1 Campo de Estudo	1
1.2 Tema e importância da pesquisa	3
1.3 Objetivos	5
1.4 Metodologia, delimitações e limitações	6
1.5 Estrutura do trabalho	9
2 REVISÃO DA LITERATURA	11
2.1 Antecedentes e fundamentos do modelo básico	11
2.2 Fronteira de produção agregada	20
2.3 Fronteira de custo	25
2.4 Fronteira de produção inversa	34
2.5 Fronteira de distância-produto	38
2.6 Fronteira determinística de produção múltipla	41
2.7 Síntese e considerações	45
3 METODOLOGIA	50
3.1 Fundamentos teóricos	51
3.1.1 Tecnologia de produção e eficiência técnica	51
3.1.2 Funções de produção de fronteira determinística e eficiência técnica	56

3.1.3	Funções de produção de fronteira estocástica e eficiência técnica	60
		vi
3.1.3.1	Introdução	60
3.1.3.2	Estimação pelo método da máxima verossimilhança	62
3.1.3.3	Estimação pelo método dos mínimos quadrados ordinários corrigido	64
3.1.3.4	Estimação pelo método da análise de correlação canônica corrigida: proposição	67
3.1.3.5	Estimação da eficiência técnica	70
3.2	Modelos de Análise	75
3.2.1	Especificações gerais	75
3.2.1.1	Delineamento básico	75
3.2.1.2	Verificação de hipóteses	78
3.2.2	Especificações particulares	82
3.2.2.1	Fronteira de produção agregada	83
3.2.2.2	Fronteira de custo	85
3.2.2.3	Fronteira de produção inversa	88
3.2.2.4	Fronteira de distância-produto	90
3.2.2.5	Fronteira de produção múltipla: proposição	92
3.3	Dados amostrais	98
3.3.1	Metodologia básica	98
3.3.2	Geração e seleção das amostras aleatórias	99
3.4	Análise comparativa	101
3.5	Programas computacionais	104
4	RESULTADOS E INTERPRETAÇÃO	106
4.1	Construção dos dados amostrais	106
4.2	Estimação e verificação dos modelos de análise	107
4.3	Análise comparativa	111
4.4	Interpretação dos resultados	113
5	CONCLUSÕES E CONSIDERAÇÕES FINAIS	118

REFERÊNCIAS BIBLIOGRÁFICAS	123
APÊNDICES	129

LISTA DE TABELAS

Tabela 4.1	Estimativas dos parâmetros das funções de fronteira estocástica, segundo cada modelo	107
Tabela 4.2	Teste da aderência de U_i pelo qui-quadrado	109
Tabela 4.3	Teste da significância global do ajustamento canônico do modelo PM	110
Tabela 4.4	Estimativas das elasticidades marginais de cada produto com relação a cada insumo, segundo o modelo PM	110
Tabela 4.5	Matriz de correlação ordinal (de Spearman) entre os índices de eficiência técnica estimados em cada modelo	112
Tabela 4.6	Matriz das somas dos postos (entre parênteses) e das diferenças das somas para cada modelo	113
Tabela 4.7	Discriminação das 15 UTD's mais eficientes e das 15 mais ineficientes, segundo cada modelo	115

LISTA DE FIGURAS

Figura 3.1	Gráfico convencional de uma isoquanta/isocusto	53
Figura 3.2	Medida de eficiência técnica de Debreu-Farrell orientada para o produto	55
Figura 4.1	Percentis das distribuições dos índices de eficiência estimados pelos modelos de análise	117

LISTA DE APÊNDICES

Apêndice A (seção 4.2)	Histograma de frequências absolutas de U_i , segundo o modelo PAI	129
Apêndice B (seção 4.2)	Histograma de frequências absolutas de U_i , segundo o modelo PAC	130
Apêndice C (seção 4.2)	Histograma de frequências absolutas de U_i , segundo o modelo CP	131
Apêndice D (seção 4.2)	Histograma de frequências absolutas de U_i , segundo o modelo PI	132
Apêndice E (seção 4.2)	Histograma de frequências absolutas de U_i , segundo o modelo DP	133
Apêndice F (seção 4.2)	Histograma de frequências absolutas de U_i , segundo o modelo PM	134
Apêndice G (seção 4.2)	Diagrama da relação entre os produtos y_{1i} e y_{2i}	135
Apêndice H (seção 4.2)	Diagrama da relação entre os insumos x_{1i} e x_{2i}	136

LISTA DE SIGLAS

ALS	modelo de fronteira estocástica de produção criado por Aigner et al. (1977)
CCA	método da análise de correlação canônica
CCCA	método da análise de correlação canônica corrigida
CD	função tipo Cobb-Douglas
CET	função com elasticidade constante de transformação
COLS	método dos mínimos quadrados ordinários corrigido
CP	modelo de fronteira estocástica de custo
CRM	modelo de regressão de cumeira canônica
DEA	método da análise envoltória de dados
DP	modelo de fronteira estocástica de distância-produto
FRONTIER	<i>software</i> de análise econométrica
GLS	método de estimação de painel com efeitos randômicos
GLS-IV	método dos mínimos quadrados generalizados para estimação de painel com variáveis instrumentais
LIMDEP	<i>software</i> de análise econométrica
MLE	método da máxima verossimilhança
MOLS	método dos mínimos quadrados ordinários modificado
OLS	método dos mínimos quadrados ordinários
PAC	modelo de fronteira estocástica de produção agregada pelos coeficientes canônicos
PAI	modelo de fronteira estocástica de produção agregada por índice
PI	modelo de fronteira estocástica de produção inversa
PM	modelo de fronteira estocástica de produção múltipla

SAS	<i>software</i> de análise matemática, estatística e econométrica
SUR	método das regressões aparentemente não relacionadas
TRANSLOG	função transcendental logarítmica
UTD	unidade tomadora de decisão
WITHIN	método de estimação de painel com efeitos fixos

LISTA DOS PRINCIPAIS SÍMBOLOS E ABREVIATURAS

Aplicável do capítulo 3 em diante. As referências entre parênteses são do primeiro registro

ET_i	índice de eficiência técnica da i -ésima UTD (equação 3.24)
\wp	vetor de parâmetros (equação 3.24)
u_i	componente residual que é a medida da ineficiência técnica no modelo de fronteira determinística (equação 3.26)
b_0	logaritmo neperiano do coeficiente de interseção da função “média” (equação 3.30)
b_k	k -ésimo parâmetro associado ao k -ésimo insumo (equação 3.30)
e_i	i -ésimo resíduo da estimação da função “média” (equação 3.30)
E	valor esperado (equação 3.31)
e_{icols}	i -ésimo resíduo da estimação por COLS da função de fronteira determinística (equação 3.32)
V_i	componente residual aleatório que indica o ruído estatístico no modelo de fronteira estocástica (equação 3.33)
U_i	componente residual que é a medida da ineficiência técnica no modelo de fronteira estocástica (equação 3.33)
m_3	estimador do terceiro momento (\mathbf{m}_3) dos resíduos da função “média” (equação 3.38)
$LogL$	função logarítmica de máxima verossimilhança (1º parágrafo da seção 3.1.3.2)
\mathbf{s}_U^2	variância não corrigida de U_i (equação 3.40)
\mathbf{s}_V^2	variância de V_i (equação 3.41)
\mathbf{g}^\bullet	razão entre a variância de U_i e a variância de \mathbf{e}_i (equação 3.41)
\mathbf{s}^2	variância de \mathbf{e}_i (equação 3.42)
\mathbf{g}	razão entre a variância não corrigida de U_i e a variância de \mathbf{e}_i (equação 3.42)
Φ	função de densidade normal cumulativa (equação 3.42)
m_2	estimador do segundo momento (\mathbf{m}_2) dos resíduos da função “média”

	(equação 3.48)
\hat{b}_{0ols}	estimador do coeficiente de interseção da função de fronteira estocástica (equação 3.51)
\hat{b}_{0ols}	estimador do coeficiente de interseção da função “média” (equação 3.51)
A	vetor dos coeficientes canônicos das variáveis referentes aos produtos (equação 3.57)
B	vetor dos coeficientes canônicos das variáveis referentes aos insumos (equação 3.57)
Cov	covariância (equação 3.59)
Var	variância (equação 3.59)
Σ	matriz de covariância (equação 3.59)
$Corr$	correlação (equação 3.60)
r	maior correlação entre as variáveis canônicas (equação 3.60)
R	matriz de correlação (equação 3.67)
RC	vetor de raízes características (equação 3.67)
\hat{b}_{0ccca}	estimador do coeficiente de interseção da função de produção múltipla de fronteira estocástica (equação 3.69)
\hat{b}_{0cca}	estimador do coeficiente de interseção da função de produção múltipla “média” (equação 3.69)
f	função densidade normal padronizada (equação 3.73)
M	moda (equação 3.74)
m	média populacional de U_i (equação 3.78)
\bar{m}	média populacional da distribuição normal truncada de U_i para dados de painel (equação 3.83)
\bar{e}_i	i -ésimo resíduo da estimação da função de fronteira estocástica para dados de painel (equação 3.83)
s_{\cdot}^2	variância populacional da distribuição normal truncada de U_i para dados de painel (equação 3.84)
\hat{s}_{\cdot}	estimador do erro-padrão da distribuição normal truncada de U_i para dados de painel (equação 3.88)

M_i^\bullet	estimador da média da distribuição normal truncada de U_i para dados de painel (equação 3.88)
LR	estatística da razão de verossimilhança para testar a existência de uma fronteira (equação 3.93)
$\sqrt{b_1}$	estatística da inclinação da distribuição de e_i para testar a existência de uma fronteira (equação 3.94)
$c_{j(\text{calc})}^2$	estatística de um teste com distribuição qui-quadrado para o j -ésimo modelo (equação 3.95)
m_4	estimador do quarto momento (m_4) dos resíduos da função “média” (conjunto de equações 3.97)
m_6	estimador do sexto momento (m_6) dos resíduos da função “média” (conjunto de equações 3.97)
a_q^{pai}	q -ésimo parâmetro associado à participação do q -ésimo produto, segundo o modelo PAI (equação 3.103)
a_q	q -ésimo parâmetro associado ao q -ésimo produto (equação 3.105)
k	parâmetro indicativo dos retornos de escala (equação 3.108)
$\frac{h_k}{k}$	k -ésimo parâmetro associado ao k -ésimo preço de insumo normalizado (equação 3.108)
b_k^{pi}	k -ésimo parâmetro associado à participação do k -ésimo insumo, segundo o modelo PI (equação 3.118)
a_q^{dp}	q -ésimo parâmetro associado ao q -ésimo produto normalizado, segundo o modelo DP (equação 3.120)
$b_{0\text{ccca}}$	coeficiente de interseção da função de produção múltipla de fronteira estocástica (equação 3.123)
$s_{\ln x_{ki}}$	desvio-padrão do k -ésimo insumo logaritmado (equação 3.129)
r_q	q -ésimo coeficiente de correlação entre as variáveis canônicas (equação 3.131)
c_{calc}^2	estatística de um teste com distribuição qui-quadrado (equação 3.131)
$c_{pm}^2(\text{n.s.})$	valor crítico do centésimo percentil da distribuição qui-quadrado com pm graus de liberdade (equação 3.131)

F_{calc}	estatística de um teste com distribuição F (equação 3.133)
R^2	coeficiente de determinação (equação 3.134)
R	coeficiente de correlação amostral entre cada variável canônica e as respectivas variáveis de produção ou de insumos, quando os dados estão padronizados (equação 3.136)
$EMg(y_q, x_k)$	elasticidade marginal do q -ésimo produto com relação ao k -ésimo insumo (equação 3.138)
$PMg(y_{qi}, x_{ki})$	produtividade marginal do q -ésimo produto com relação ao k -ésimo insumo para a i -ésima UTD (equação 3.139)
$TMgTP(y_{qi}, y_{q'i})$	taxa marginal de transformação entre o q -ésimo produto e o q' -ésimo produto para a i -ésima UTD (equação 3.141)
$TMgSI(x_{ki}, x_{k'i})$	taxa marginal de substituição entre o k -ésimo insumo e o k' -ésimo insumo para a i -ésima UTD (equação 3.142)
$\text{Conv}(x_{1i}, \dots, x_{mi})$	condição de convexidade da isoquanta da função de produção múltipla (equação 3.143)
$\text{Conv}(y_{1i}, \dots, y_{mi})$	condição de convexidade da curva de transformação da função de produção múltipla (equação 3.144)
h	titulação conjunta das funções de transformação e de produção (item III, seção 3.3.2)
$SP_{j,j'}$	coeficiente de correlação ordinal de Spearman entre o j -ésimo e o j' -ésimo modelo (equação 3.151)
$t_{j,j'}(\text{calc})$	estatística de um teste com distribuição t envolvendo o j -ésimo e o j' -ésimo modelo (equação 3.152)
FF	estatística de um teste de Friedman com distribuição F (equação 3.154)
DSM	diferença significativa mínima para o teste de Friedman com comparações múltiplas (equação 3.156)

RESUMO

Desde o aparecimento do modelo de produção de fronteira estocástica, em 1977, tem crescido o interesse pela análise econométrica da eficiência produtiva. Todavia, este tipo de análise é considerado problemático com relação ao enfoque de múltiplos produtos. O tratamento multiequacional não está suficientemente desenvolvido e a abordagem uniequacional depende de condições especiais, algumas vezes não alcançadas. Esta pesquisa trata exclusivamente dos modelos de fronteira estocástica a uma equação, direcionados à estimação de índices relativos de eficiência técnica para o caso de múltiplos insumos e múltiplos produtos. A literatura referente a estes modelos é integrada pelas fronteiras de produção agregada, de produção inversa, de distância e de valor dual. Estas fronteiras têm sido estimadas separadamente e sem muita preocupação em testar as especificações adotadas. Neste contexto, a pesquisa propõe uma alternativa metodológica em que o processo de estimação considera cada um dos vários produtos como uma variável dependente. Esta função de produção múltipla é estimada por análise de correlação canônica, usando dados simulados de *cross-section* sobre dois produtos e dois insumos, gerados segundo uma tecnologia com elasticidades constantes de transformação e de substituição. A correspondente fronteira estocástica é quantificada por meio da correção do coeficiente de interseção oriundo daquela estimação. Este novo procedimento é batizado de “análise de correlação canônica corrigida”. À exceção dos escores de eficiência técnica estimados pelos modelos de fronteira de produção agregada por um índice e de fronteira de produção inversa, os demais índices produzidos pelos outros modelos de fronteira mostram-se originários da mesma população e tem correlações ordinais estatisticamente aceitáveis entre si. Conclui-se que o novo modelo é viável como alternativa ao uso dos modelos de fronteira de produção agregada pelos coeficientes canônicos, de fronteira de custo e de fronteira de distância-produto. O trabalho também chama a atenção para o delineamento de uma importante linha de pesquisa, na qual sobressai o desenvolvimento de algoritmo que produza estimativas de máxima verossimilhança para todos os parâmetros do modelo de fronteira estocástica de produção múltipla.

ABSTRACT

Since the stochastic frontier production model was born, in 1977, interest on productive efficiency econometric analysis has been increasing. But this kind of analysis is troublesome with regard to multiple outputs. The multiple-equation approach is not well developed and single-equation approach is subject to special conditions sometimes not reached. This research deals only with single-equation stochastic frontier models with the purpose of estimating technical efficiency relative indices. The literature in this area presents aggregated production, inverse production, distance and value dual frontier models. These frontiers have been estimated separately without much concern with misspecification tests. In this research, a methodological alternative in which the estimation process treats each one of multiple outputs as a dependent variable is proposed. The approach is illustrated with a set of two outputs and two inputs cross-section simulated data generated by a constant elasticity transformation and substitution production function. The correspondent stochastic frontier is estimated by canonical correlation analysis with correction of the intercept. This procedure is new. Except for technical efficiency scores estimated through aggregate production by index and inverse production frontier models, every other frontier model indices comes from the same population and have ordinal correlations statistically acceptable. We conclude that the new model is a feasible alternative method for estimating aggregated production frontier by canonical coefficients, cost frontier and output distance frontier models. An important research line opened by this work deals with the construction of an algorithm for finding maximum likelihood estimates of all parameters for the multiple production stochastic frontier model.

INTRODUÇÃO

1.1 CAMPO DE ESTUDO

Em termos gerais, o campo de estudo da presente pesquisa é concernente com a medida e a explicação da performance do produtor.

É muito comum nas análises sobre o desempenho de empresas públicas e privadas produtoras de bens e serviços, aqui também chamadas de unidades tomadoras de decisão (UTD's), descrevê-las como sendo mais ou menos “eficientes”, ou mais ou menos “produtivas”. Estes dois conceitos, embora intimamente relacionados, representam hipóteses distintas com relação aos determinantes da performance do produtor, quando ela é medida pela eficiência e pela produtividade.

A produtividade de uma unidade de produção é frequentemente entendida como sendo a razão entre o valor ou quantidade de seus produtos e o valor ou quantidade de seus insumos, sendo bem fácil de computar se a UTD usa um único insumo para produzir um único produto. Como a grande maioria das unidades utiliza vários insumos para produzir vários produtos, deve-se agregar os primeiros no denominador e os últimos no numerador, de tal modo que a produtividade continue sendo o quociente entre duas grandezas escalares. Este tipo de produtividade varia devido a diferenças na tecnologia de produção, a diferenças na eficiência do processo produtivo e a diferenças no ambiente em que ocorre a produção. O componente da produtividade, que decorre de diferenças na eficiência do processo produtivo, é frequentemente denominado de “eficiência produtiva”.

A eficiência produtiva de uma UTD, por sua vez, é medida através da comparação entre os valores observados e os valores ótimos de seus produtos e insumos. Esta comparação pode tomar a forma da razão entre a produção observada e a produção potencial máxima alcançável, dado um conjunto de insumos, ou a forma da razão entre o montante potencial mínimo de insumo requerível e o montante efetivamente empregado, dado um conjunto de produtos, ou alguma combinação destas duas razões. Nestas duas comparações, o valor potencial ou ótimo é definido em termos de possibilidades de

produção, e a medida de eficiência é eminentemente técnica (eficiência técnica). Quando o ótimo é definido em termos de custos, receitas, lucros, ou qualquer outro objetivo da UTD sujeito às restrições de quantidades e de preços, a eficiência em questão é a econômica (eficiência alocativa).

O processo de mensuração destas medidas de performance do produtor, via envelopamento dos dados limitados por regiões de máxima produção ou de mínimo custo (fronteira), envolve uma boa variedade de técnicas econométricas e de programação matemática, neste último caso englobadas sob a denominação de análise envoltória de dados (*data envelopment analysis* - DEA).

O método econométrico convencional, ao contrário do enfoque DEA, tem precisado experimentar substanciais modificações para atender aos novos requisitos deste tipo de análise. Desde o marco metodológico fixado por Cobb & Douglas (1928), dando início à estimação empírica de funções de produção, decorrem cerca de 40 anos até a análise econométrica da eficiência produtiva ser levada a efeito. Daquela data até os anos 50, as funções de produção são largamente usadas para estudar a distribuição funcional de renda entre os fatores capital e trabalho, a nível macroeconômico. Na década de 50, desenvolve-se a análise empírica de estruturas microeconômicas de produção, porém centrada nos custos, e não na produção propriamente dita.

A análise de eficiência, sob a ótica de funções de fronteira, é delineada nas pesquisas de Debreu (1951) e Koopmans (1951) e, principalmente, Farrell (1957). Este último define a eficiência técnica em termos de desvios observados até uma idealizada isoquanta de fronteira. Tal enfoque cai naturalmente em um tratamento em que a ineficiência é expressada pelos resíduos de um modelo de regressão. Não obstante, somente 11 anos após, a análise formal de modelos deste tipo é registrada na literatura. Até então, todo o esforço empírico relacionado ao estudo da produção e do custo desenvolve-se independentemente da ótica de fronteiras de produção. É a necessidade de operacionalizar esta nova abordagem que dá início a uma série de importantes avanços por que passam as técnicas econométricas tradicionais.

Estes aperfeiçoamentos começam com a introdução dos modelos de fronteira determinística, em 1968, evoluem para a fronteira estocástica, em 1977, e experimentam substancial especialização com a abordagem de painel de dados, surgida em 1984. No entanto, até fins da década de 80, o emprego de modelos uniequacionais mantém-se restrito ao uso de fronteiras de produção agregada, em que os vários produtos são juntados em uma única variável dependente. A partir de então, a análise econométrica de eficiência incorpora melhoramentos expressivos com a utilização das fronteiras de valor dual, de produção inversa e de distância.

Todos esses modelos de fronteira, porém, não utilizam métodos de estimação direta de uma função de produção, em que pelo menos os produtos principais sejam representados por variáveis desagregadas. Este tipo de modelo é denominado, nesta pesquisa, de “fronteira de produção múltipla”. Seu enfoque é bastante recente, cabendo a Ruggiero (1996) a primeira contribuição, através de um modelo híbrido DEA/econométrico com fronteira determinística de produção do tipo Cobb-Douglas.

Dentro deste contexto, a presente pesquisa trata, pioneiramente, da especificação e estimação de um modelo uniequacional de fronteira estocástica de produção múltipla, voltado para a mensuração da eficiência técnica relativa.

1.2 TEMA E IMPORTÂNCIA DA PESQUISA

Segundo Lovell (1995) as fronteiras de custo, de receita e de lucro acomodam múltiplos produtos na mensuração econométrica uniequacional de eficiência, porém, a fronteira de produção não. Para o referido pesquisador, existem três opções à disposição. Uma, é agregar os produtos em um índice e estimar uma fronteira de produção agregada. Outra, é agregar os insumos em um índice e estimar uma fronteira de necessidades de insumo (ou de produção inversa). E, por fim, ele indica a estimação de uma fronteira de distância-produto.

Apenas pela designação das referidas alternativas pode-se deduzir que elas não envolvem a estimação direta de uma função de produção múltipla, na qual os produtos

permaneçam desagregados em função dos insumos também desagregados. O problema de estimar uma equação de regressão, quando existem duas ou mais variáveis dependentes, é substancialmente resolvido por Vinod (1968) por meio da técnica multivariada de análise de correlação canônica (*canonical correlation analysis* - CCA). Através da referida técnica é possível estimar, diretamente, uma função de produção múltipla, como uma função implícita de todos os produtos e insumos.

A despeito do extraordinário progresso verificado na análise econométrica de eficiência produtiva, passam-se 28 anos, desde a introdução dos modelos de fronteira determinística (Aigner & Chu, 1968), até ocorrer a primeira estimação direta de uma função de fronteira determinística de produção múltipla (Ruggiero, 1996). De lá para cá, o enfoque de fronteira de produção múltipla tem sido pouco explorado, apesar da crescente aplicação da análise econométrica de eficiência em ambientes com tecnologia multi-produto.

Por conseguinte, não é sem motivo que os especialistas têm sido unânimes em admitir que uma das vantagens do método DEA sobre a econometria é a habilidade de determinar escores de eficiência relativa para tecnologias com múltiplos insumos e múltiplos produtos. Nessa direção, Kumbhakar (1996) discute a modelagem da ineficiência técnica e da ineficiência alocativa, sob suposições de minimização de custos e de maximização de lucros, dando especial ênfase ao caso de múltiplos insumos e de múltiplos produtos. Contudo, ao avaliar este esforço de Kumbhakar e os de outros pesquisadores, Lovell (1996) exprime o seu desânimo pelo fato de o enfoque econométrico não ter evoluído além da estimação de fronteiras de valor dual e de distância.

As conseqüências dessas constatações podem ser o comprometimento da análise econométrica de eficiência em algumas situações. A utilização de uma fronteira de produção agregada não é adequada, quando existe produção conjunta dos múltiplos produtos em um único processo produtivo. Sobre este fenômeno, Karlson (1986) cita, como exemplo, o caso de uma UTD produtora apenas de dois tipos de serviços de transporte: passageiros e cargas. Para essa unidade, os custos marginais relativos dos serviços vendidos para diferentes classes de consumidores, dependem da combinação

tanto dos dois produtos quanto dos dois insumos usados na produção. Por isso, argumenta Karlson, é impossível construir algum produto agregado homogêneo, chamado “serviço de transporte”, para ser vendido aos dois tipos de consumidores.

Outra adversidade, é a inexistência de um insumo dominante para estimar a fronteira de produção inversa. Para a fronteira de distância -produto, a dificuldade pode residir no não-cumprimento do requisito de homogeneidade. Já o uso de fronteiras de valor dual, para mensurar unicamente a eficiência técnica, envolve condições muito especiais. Na ausência destas condições, tais fronteiras podem ser não duais ou estimar a eficiência técnica e a eficiência alocativa, conjuntamente. Neste último caso, a decomposição da eficiência produtiva implica no uso de sistemas de equações e na convivência com problemas de estimação ainda não suficientemente resolvidos. Outro impedimento ao uso de fronteiras duais de custo, de lucro e de receita, pode ser a falta de dados sobre preços.

Ante as dificuldades enunciadas nos dois últimos parágrafos, é possível que uma determinada análise de eficiência, se operacionalizada com rigor científico, fique inviabilizada diante de tantas condições a satisfazer. Inserido nessa problemática, o novo modelo, proposto nesta pesquisa, tem sua importância sensivelmente aumentada, pois pode tornar-se a única alternativa de aplicação empírica, em alguns casos.

1.3 OBJETIVOS

O objetivo geral desta pesquisa é desenvolver e apresentar uma metodologia alternativa para estimar diretamente a eficiência técnica de UTD's com tecnologia de múltiplos produtos, por meio de um modelo econométrico uniequacional de fronteira estocástica de produção múltipla. Na busca deste intento, vislumbra-se, especificamente, alcançar as seguintes realizações:

- 1) Conhecer as condições e requisitos estatísticos, econométricos e de teoria econômica previstos na literatura teórico-aplicada pertinente;

- 2) Verificar a possibilidade de estender à análise CCA os procedimentos de estimação de fronteira estocástica usados no método dos mínimos quadrados ordinários corrigido (*corrected ordinary least squares* - COLS);
- 3) Construir uma amostra aleatória de dados que viabilize a especificação, a estimação e a verificação dos diferentes modelos de análise;
- 4) Elaborar programas computacionais para processar os modelos de análise com os dados das amostras aleatórias geradas;
- 5) Avaliar a consistência entre os índices estimados pelo modelo proposto e os índices estimados pelos modelos existentes na literatura e
- 6) Identificar necessidades e possibilidades para a realização de pesquisas futuras.

1.4 METODOLOGIA, DELIMITAÇÕES E LIMITAÇÕES

A estrutura metodológica do presente trabalho é integrada por um conjunto de fundamentos teóricos, métodos, técnicas e modelos de estimação, de verificação e de análise que dizem respeito aos requisitos estatísticos, econométricos e de teoria econômica envolvidos com a aplicação de modelos uniequacionais na análise econométrica de eficiência técnica. Os referidos modelos de fronteira estocástica são também conhecidos como de “resíduos compostos”, porque seus resíduos são desmembrados em um componente bicaudal (normal) e um unicaudal (não-normal) indicativo da eficiência técnica. Os dados para a análise empírica são construídos através de um gerador de variáveis aleatórias e constituem uma amostra com dois produtos e dois insumos.

A partir da estimação e verificação dos modelos vigentes na literatura e do modelo proposto, utiliza-se os resíduos compostos para mensurar os índices de eficiência técnica para cada UTD, segundo cada modelo de fronteira estimado. Na etapa seguinte, são utilizadas técnicas estatísticas não-paramétricas para avaliar a consistência dos índices estimados, principalmente em comparação com os fornecidos pelo novo modelo. Todo o trabalho de geração da base de dados, de estimação, de verificação e de avaliação comparativa é realizado através do *software* denominado, abreviadamente, LIMDEP

(*limited dependent variable models* - Greene, 1995), e em planilhas do Excel. A utilização de um bom número de técnicas não triviais requer grande esforço na programação de alguns algoritmos, nos quais se utilizam comandos e funções disponíveis no LIMDEP.

A estratégia adotada no detalhamento da metodologia leva em conta que o objetivo primordial da pesquisa é desenvolver e apresentar um novo modelo, sob condições bem definidas. Por isso, na tarefa de especificação metodológica considera-se, preferencialmente, os procedimentos referentes aos requisitos exigidos para assegurar condições satisfatórias de comparabilidade dos resultados.

Seguindo esse direcionamento, utiliza-se a classificação de Lovell (1993 e 1995) para identificar os modelos econômicos uniequacionais usualmente empregados no caso de tecnologias multi-produto. Eles são as fronteiras de valor dual, de produção agregada, de produção inversa e de distância-produto. Considera-se todos esses modelos a uma equação, porque os modelos multiequacionais são ainda muito incipientes (Lovell, 1993) e encerram extraordinárias dificuldades de estimação (Greene, 1993b).

Quanto à modalidade de fronteira, adota-se a do tipo estocástico, já que “por motivos práticos, os econometristas têm abandonado a fronteira determinística ...” (Greene, 1993b). A fronteira estocástica requer a especificação de uma distribuição para os resíduos indicativos da eficiência técnica. Como os estimadores COLS empregados na pesquisa, até o momento, possuem propriedades amostrais finitas conhecidas apenas para resíduos seminormais, assume-se este tipo de distribuição em todos os modelos de análise.

Quando se usa dados de painel, não é necessário definir a distribuição dos resíduos unicaudais. Contudo, é preciso definir a distribuição da evolução da eficiência ao longo do tempo, uma vez que a suposição de sua invariabilidade é insustentável, especialmente quando o número de observações da série temporal aumenta. Além deste inconveniente, o uso do painel de dados, na presente pesquisa, é proibitivo, porque ainda não se conhecem as propriedades amostrais finitas dos seus estimadores COLS. Conseqüente mente, a pesquisa contempla, exclusivamente, o emprego de dados de *cross-section*.

No que concerne à forma funcional, segue-se a tendência observada nas aplicações empíricas consultadas. Na sua grande maioria, elas adotam a especificação de uma tecnologia de produção Cobb-Douglas.

A estimação dos modelos uniequacionais de fronteira estocástica existentes na literatura não é feita por máxima verossimilhança (MLE), porque este método ainda não está desenvolvido para estimar, direta e conjuntamente, os parâmetros de funções de fronteira com mais de uma variável dependente. O método COLS também não, porém ele pode ser estendido à análise de correlação canônica, que é capaz de realizar a referida estimação. Assim, para padronizar os procedimentos, estima-se os coeficientes de regressão dos modelos vigentes, inicialmente, por mínimos quadrados ordinários (OLS). A seguir, por COLS, estima-se as fronteiras estocásticas correspondentes, por intermédio da correção dos coeficientes de interseção obtidos por OLS. Dos resíduos daí resultantes, extraí-se as estimativas do componente unicaudal indicativo da eficiência técnica, aplicando a formulação criada por Battese & Coelli (1988).

Por sua vez, a função de produção múltipla é estimada, inicialmente, por CCA, transformando-se em função de fronteira pela correção do coeficiente de interseção. A essência desta nova abordagem é a incorporação do vetor de resíduos compostos do modelo de fronteira estocástica. Esta incorporação é exequível, pois os resíduos de regressão canônica provêm de modelos com propriedades estatísticas análogas às dos modelos que geram resíduos de mínimos quadrados ordinários. Assim, para estimar a fronteira estocástica de produção múltipla, aplicam-se os mesmos procedimentos do método COLS, porém, num contexto de correlação canônica. Este procedimento é batizado, nesta pesquisa, de “análise de correlação canônica corrigida” (CCCA). Os resíduos oriundos da aplicação do método CCCA são, então, usados para estimar os índices de eficiência técnica através da mesma fórmula empregada nos demais modelos.

Uma limitação do método CCA é a inexistência de desenvolvimento teórico para os erros-padrão dos coeficientes canônicos. Por este motivo, recorre-se ao método de reamostragem conhecido como *jackknife*. O erro-padrão do coeficiente de interseção assim estimado é utilizado no cálculo do erro-padrão do coeficiente de interseção

corrigido. Esta estimativa e as dos erros-padrão dos demais coeficientes canônicos são empregadas no cálculo das estatísticas t assintóticas necessárias aos testes de significância.

Para a função de produção múltipla, também é testada a significância global do ajustamento canônico e é verificada a magnitude da proporção da variância amostral que é explicada pelas variáveis canônicas. Por tratar-se de uma nova abordagem, investiga-se, também, algumas implicações econômicas relacionadas às condições de primeira e de segunda ordens.

Finalmente, em sintonia com a motivação central da pesquisa, procede-se à análise comparativa dos resultados, sem precisar ir além da investigação sobre a correlação e a variabilidade dos índices estimados pelos diferentes modelos de análise.

1.5 ESTRUTURA DO TRABALHO

O capítulo 2 sintetiza os principais aspectos levantados na revisão da literatura pertinente ao tema da pesquisa. Procura-se chegar ao estado da arte na questão da mensuração da eficiência técnica relativa através da estimação de modelos econométricos uniequacionais de fronteira estocástica. Após classificados, os referidos modelos são descritos por categoria, enfocando-se as implicações metodológicas e empíricas de cada um. Inclui-se, neste capítulo, as únicas referências encontradas sobre a estimação direta de uma fronteira de produção múltipla, ainda que tais fronteiras sejam do tipo determinístico.

No capítulo 3 são detalhados os fundamentos teóricos relevantes quanto à representação da tecnologia de produção e à mensuração do grau de eficiência com que a referida tecnologia é operada. Esta formalização passa pela explicitação das funções de produção de fronteira estocástica e pelos métodos disponíveis para estimá-las, desembocando na base teórica inerente à estimação dos índices de eficiência técnica para cada UTD, individualmente. É neste capítulo em que se definem todas as especificações estatísticas, econométricas e de teoria econômica necessárias à aplicação empírica dos modelos de análise. Este capítulo também apresenta a estratégia metodológica seguida na

construção dos dados amostrais, na realização da análise comparativa dos resultados e na elaboração dos programas de computador.

No capítulo 4 concentram-se os resultados da utilização dos programas computacionais elaborados de acordo com a metodologia desenvolvida no capítulo anterior. A tarefa de quantificação, basicamente, consta da construção da base de dados, da estimação e verificação dos modelos de análise e do processamento das técnicas não-paramétricas utilizadas na comparação dos índices de eficiência estimados. A interpretação dos resultados produzidos nessas fases é a sustentação para as principais conclusões e considerações finais emitidas no capítulo seguinte.

A documentação do trabalho completa-se com os apêndices contendo os histogramas dos índices de eficiência estimados em cada modelo de análise e os diagramas com os dados sobre os dois produtos e os dois insumos gerados por simulação.

REVISÃO DA LITERATURA

Este capítulo descreve os principais aspectos metodológicos referentes à especificação e estimação de modelos econométricos uniequacionais para análise de eficiência técnica, começando pelo modelo básico de fronteira estocástica de produção. A descrição dos modelos subsequentes cobre exclusivamente os que enfocam o contexto de múltiplos produtos, segundo a classificação de Lovell (1995).

Na penúltima seção é abordada a estimação direta de funções de fronteira de produção múltipla, sob o enfoque determinístico. A especificação e estimação deste tipo

de função, no caso de fronteiras estocásticas, constitui a essência do modelo proposto na pesquisa. Ele é desenvolvido e apresentado no capítulo 3.

Na última seção é apresentada uma síntese crítica dos aspectos e especificações mais relevantes para a uniformização de procedimentos metodológicos.

2.1 ANTECEDENTES E FUNDAMENTOS DO MODELO BÁSICO

Pelo enfoque econométrico tradicional, quando se estima uma função de custo ou de produção, obtêm-se estimativas da função “média” de produção ou de custo que melhor ajusta os dados observados. Isto é inconsistente com a concepção teórica de custo ou de produção, pela qual a função delinea uma região de custo mínimo para cada combinação de níveis de produtos, ou de máximo produto para cada combinação de níveis de insumos.

Farrell (1957) é um dos primeiros a propor a operacionalização empírica da idéia de custo mínimo. Ele idealiza uma tecnologia de fronteira caracterizada por uma única e eficiente isoquanta. Assume uma função de produção com retornos constantes de escala e decompõe a eficiência produtiva em eficiência técnica e eficiência alocativa. Assim, interpreta o uso excessivo de insumos como tecnicamente ineficiente porque é possível reduzir todos os insumos, proporcionalmente, sem reduzir o nível de produto: neste caso a firma está operando acima da isoquanta unitária. A ineficiência alocativa, por sua vez, refere-se a incorretas proporções no uso dos fatores: a firma está no lugar errado na isoquanta. E o grau de ineficiência do custo é medido pela razão da distância da linha de isocusto original com relação a linha de isocusto que tangencia o custo mínimo.

A isoquanta unitária eficiente é naturalmente não-observada. Ela deve ser estimada através de uma amostra de observações sobre produtos, insumos, preços e custos de firmas que podem ser eficientes ou ineficientes. Farrell propõe desenhar a fronteira seguindo uma linha convexa de livre descarte das razões de insumo-produto usando programação linear. Este procedimento, mais tarde, torna-se conhecido como DEA. Ele é

não-paramétrico, pois não é baseado em nenhum modelo explícito de fronteira ou de relação das observações na fronteira, as quais não podem ficar acima da mesma.

Não obstante, Farrell, quase que numa reflexão tardia, propõe uma envoltória convexa paramétrica do tipo Cobb-Douglas. Ele reconhece a indesejabilidade de impor uma forma funcional específica para a fronteira, mas, também, observa a vantagem de expressar a fronteira de uma forma matemática simples. Infelizmente, Farrel não executa a sua própria sugestão.

Uma década após, Aigner & Chu (1968) são os primeiros a implementar a sugestão de Farrell. Eles consideram somente os efeitos dos choques aleatórios sobre o processo de produção, o que faz com que todas as diferenças de eficiência técnica fiquem incluídas no termo residual. Assumem que os erros de medida em todas as variáveis são negligenciáveis e, por simplicidade, especificam o seguinte modelo de produção Cobb-Douglas, a um produto x_0 e a dois insumos x_1 e x_2 :

$$x_0 = A x_1^a x_2^b u \quad (2.1)$$

onde: u é o termo aleatório e

A, a, b são os parâmetros a estimar.

Definem, então, como problema central, obter a função estimada

$$\hat{A} \hat{x}_1^{\hat{a}} \hat{x}_2^{\hat{b}} = \hat{x}_0 \quad (2.2)$$

$$\text{tal que} \quad \hat{A} \hat{x}_1^{\hat{a}} \hat{x}_2^{\hat{b}} \geq x_0 \quad (2.3)$$

Tomando os logaritmos de ambos os lados de (2.1), eles reescrevem a referida equação como

$$X_0 = XC + e \quad (2.4)$$

$$\text{e} \quad XC \geq X_o \quad (2.5)$$

onde: $X_0 = \log x_0, \quad X_1 = \log x_1, \quad X_2 = \log x_2;$

$X = [1 \quad X_1 \quad X_2];$

$C = [A \quad a \quad b]'$ e

e é o vetor dos resíduos medidos.

Aigner e Chu seguem dois modos de enfocar os problemas de estimação das equações (2.4) e (2.5). Num deles, minimizam a soma dos quadrados dos resíduos $e'e$ sujeita a (2.5), encontrando o vetor C que minimiza a função quadrática

$$e'e = (X\hat{C} - X_0)'(X\hat{C} - X_0) = \hat{C}'X'X\hat{C} - 2\hat{C}'X'X_0 + X_0'X_0 \quad (2.6)$$

sujeita a $X\hat{C} \geq X_0$

No outro, utilizam programação linear, dado que os choques residem somente em um lado da fronteira de produção. Minimizam a soma dos resíduos como uma função linear de perda, ou seja,

$$\min \quad l'e = l'(X\hat{C} - X_0), \quad l' = [1 \ 1 \ . \ . \ 1] \quad (2.7)$$

sujeita a $X\hat{C} \geq X_0$

e $\hat{C} \geq 0$

O principal problema destes enfoques é que eles são extremamente sensíveis aos *outliers*. Para resolver isto, Aigner e Chu sugerem o descarte de algumas observações, idéia que é implementada por Timmer (1971) ao desenvolver as fronteiras probabilísticas. Estas fronteiras também são determinadas por programação matemática, porém, alguma especificada proporção de observações pode localizar-se acima da fronteira. A seleção desta proporção é essencialmente arbitrária, carecendo de justificação econômica ou estatística. Não obstante, permanecia sem solução a necessidade de reconciliar as observações acima da fronteira com o conceito de fronteira como um nível máximo de produção. Tipicamente, isto é efetuado recorrendo-se ao erro de medida das observações extremas. Deste modo, parecia naturalmente preferível incorporar a possibilidade de erros de medida e de outros choques não observáveis, de uma maneira menos arbitrária.

Entretanto, as estimativas produzidas pela programação matemática não têm propriedades estatísticas conhecidas (erros-padrão, razões t , etc.). Assim, na falta de suposições estatísticas sobre C ou e em (2.4), os resultados inferenciais não podiam ser obtidos.

Na tentativa de conseguir alguma base estatística, Schmidt (1976) adiciona explicitamente um resíduo unicaudal na função (2.1), representando-a, de forma geral, por

$$y_i = f(x_i; \mathbf{b}) + \mathbf{e}_i, \quad i = 1, \dots, N \quad (2.8)$$

onde: y_i é o produto escalar máximo possível de obter com o vetor x_i (não estocástico) de insumos;

\mathbf{b} é um vetor desconhecido de parâmetros a estimar e

$\mathbf{e}_i \leq 0$.

Schmidt constata que as estimativas de Aigner e Chu, obtidas por programação matemática, são, também, de máxima verossimilhança, sob especificações particulares das distribuições dos erros. Mais especificamente, se \mathbf{e}_i de (2.8) ou u de (2.1) tem distribuição exponencial, a programação linear produz estimadores de máxima verossimilhança, enquanto que a programação quadrática só produz estimadores de máxima verossimilhança se u (ou \mathbf{e}_i) tem distribuição seminormal.

Apesar do esforço dispendido por Schmidt, suas constatações tem pouco sentido prático, em virtude da violação das usuais condições de regularidade para a aplicação da máxima verossimilhança. Ou seja, uma vez que $y_i \leq f(x_i; \mathbf{b})$, o campo de variação da variável dependente y depende dos parâmetros a serem estimados, de forma que os teoremas usuais não podem ser invocados para determinar as distribuições assintóticas das estimativas dos parâmetros. Nestas circunstâncias, não fica claro o quanto se conhecia a respeito da fronteira, depois dela ter sido estimada.

A primeira tentativa de tornar o modelo de Aigner e Chu acessível à análise estatística, deve-se a Afriat *apud* FORSUND et al. (1980), que introduz algumas suposições e reescreve-o como

$$y = f(x)e^{-u} \quad (2.9)$$

$$\text{ou} \quad \ln y = \ln[f(x)] - u \quad (2.10)$$

onde: y é um dado produto escalar;

x é um vetor de variáveis-insumo exógenas (independentes de u);

$\ln[f(x)]$ é linear no caso Cobb-Douglas exibido em (2.4) e

$u \geq 0$ é independente e idênticamente distribuído (iid) no intervalo

$$0 \leq e^{-u} \leq 1.$$

Afriat assume uma distribuição beta com dois parâmetros para e^{-u} e propõe a estimação do modelo por MLE. Este procedimento, segundo Richmond (1974), resulta em uma distribuição gama para u .

Dois anos após, Aigner et al. (1976) constroem uma estrutura de erro mais razoável do que a puramente unicaudal. Especificamente, eles assumem que

$$\mathbf{e}_i = \begin{cases} \mathbf{e}_i^* / \sqrt{(1-q)}, & \text{se } \mathbf{e}_i^* > 0, \\ \mathbf{e}_i^* / \sqrt{q}, & \text{se } \mathbf{e}_i^* \leq 0 \end{cases} \quad i = 1, \dots, N, \quad (2.11)$$

onde: \mathbf{e}_i^* (erros) são variáveis aleatórias independentes e normalmente distribuídas,

com média zero e variância σ^2 para $0 < q < 1$ e

\mathbf{e}_i^* tem distribuição normal truncada negativa ou positiva quando $q = 1$

ou $q = 0$, respectivamente.

Os idealizadores desta especificação dos erros justificam-na com o argumento de que as firmas diferem na produção de y , para um dado conjunto de insumos, de acordo com a variação aleatória na sua habilidade de utilizar a melhor tecnologia, que é uma fonte de erro unicaudal ($\mathbf{e}_i \leq 0$), e/ou devido à variação aleatória na quantidade de insumos ou erro de medida em y , que é um erro simétrico. Assim, q é interpretado como a medida de variabilidade relativa destas duas fontes de erro. Quando seu valor é igual à unidade, ele circunscreve toda a fronteira, e quando é igual a 1/2 a função “média”.

Uma contribuição primária desta estrutura de erros para a literatura é que ela permite a estimação conjunta da função a ser ajustada e do parâmetro q . Contudo, a interpretação de q , como uma medida da variabilidade relativa das fontes de erro, não está explícita na formulação da equipe de Aigner. Ainda falta modelar o processo de erros segundo esta concepção.

Outro aspecto evolutivo marcante é o fato de que todas as fronteiras empregadas, até então, são do tipo determinístico. Nesta modalidade de fronteira, todas as firmas compartilham uma família comum de fronteiras de produção, custo e lucro, e toda a variação no desempenho da firma é atribuída apenas à variação na sua eficiência com relação às referidas fronteiras. Esta concepção ignora a possibilidade real de a eficiência

ser também afetada por fatores externos inteiramente fora do controle da firma. E é, até intuitivamente, um tanto questionável juntar, em um único termo residual unicaudal, os efeitos de choques exógenos com os efeitos dos erros de medida e da própria ineficiência, e chamar tudo de “eficiência” ou “ineficiência” da firma.

Esta constatação fica bastante reforçada quando se considera, também, o ruído estatístico presente em todas as relações empíricas. Tais ruídos podem provir de erros de medida, principalmente da variável dependente, e da inadequada e incompleta especificação da equação, seja pela omissão de variáveis importantes, seja pela inadequação da forma funcional. Por isso, estas constatações valem, não só para funções de produção, como para qualquer outra equação e é, no mínimo dúvida, não distinguir o que é ruído estatístico e o que é ineficiência ou, então, assumir que o ruído é unicaudal.

É nestes argumentos em que se apoiam as fronteiras estocásticas de produção, criadas independente, mas, simultaneamente, por pesquisadores de três continentes: Aigner et al. (1977), Meeusen & van den Broeck (1977) e Battese & Corra (1977). De lá para cá, a formulação mais referenciada e seguida é a do primeiro grupo. Por isso, apenas esta abordagem é incluída nesta revisão. Ela é, por muitos, para fins de simplificação, denominada ALS.

Aigner et al. (1977) idealizam o modelo (2.8) com uma estrutura de erro dada por

$$\mathbf{e}_i = v_i + u_i \quad i = 1, \dots, N \quad (2.12)$$

onde: v_i representa o resíduo simétrico que é suposto ser iid como uma $N(0, \mathbf{s}_v^2)$ e

u_i é suposto ser distribuído independentemente de v_i , satisfazendo a condição

$$u_i \leq 0.$$

Eles consideram, particularmente, o caso em que u_i é derivado de uma distribuição $N(0, \mathbf{s}_u^2)$ truncada acima de zero, mas, também, abordam, brevemente, o caso em que u_i tem distribuição exponencial.

A lógica econômica desta especificação, segundo seus inventores, é que o processo de produção está sujeito a dois tipos de distúrbio aleatório, economicamente distinguíveis e com diferentes características. O termo não-positivo u_i reflete o fato do

nível de produto de cada firma poder estar sobre ou abaixo do seu nível na fronteira $[f(x_i; \mathbf{b}) + v_i]$. Qualquer desvio é o resultado de fatores sob o controle da firma, tais como ineficiência técnica e econômica, a vontade e o esforço do produtor e de seus empregados e, também, pode estar refletindo a ocorrência de produtos defeituosos e estragados. A fronteira pode variar aleatoriamente entre as firmas ou ao longo do tempo para a mesma firma. Por isso, ela é estocástica com o distúrbio aleatório $v_i \stackrel{>}{=} 0$ representando o resultado de acontecimentos externos à firma, favoráveis a ela ou não, tais como sorte, clima, topografia e desempenho das máquinas, além de expressar a ocorrência de erros de observação e de medida da variável y .

Um importante subproduto deste enfoque é que se pode estimar as variâncias de v_i e de u_i e, assim, avaliar as suas magnitudes relativas. Outra implicação, ainda mais importante, é que se pode medir a eficiência produtiva da firma pela razão

$$\frac{y_i}{[f(x_i; \mathbf{b}) + v_i]} \quad (2.13)$$

ao invés de
$$\frac{y_i}{[f(x_i; \mathbf{b})]} \quad (2.14)$$

Este procedimento distingue a ineficiência produtiva de outras fontes de perturbação que estão além do controle da firma. Para exemplificar melhor, Aigner e sua equipe mencionam a situação de um agricultor cuja colheita é dizimada pela seca ou pela tempestade. Por (2.13), a perda é considerada falta de sorte e, por (2.14), ineficiência do produtor.

Para fins de desenvolver os procedimentos referentes a estimação, os referidos pesquisadores consideram um modelo linear escrito, na forma matricial, como

$$y = X\mathbf{b} + \mathbf{e} \quad (2.15)$$

onde $\mathbf{e} = v + u$.

Atribuindo a Weinstein (1964) a autoria da primeira derivação da função de distribuição da soma de uma variável aleatória normal simétrica com uma variável aleatória normal truncada, eles argumentam que a derivação da função densidade de \mathbf{e} é direta, não a incluem no texto e apresentam o resultado como sendo

$$f(\mathbf{e}) = \frac{2}{\mathbf{s}} f^* \left(\frac{\mathbf{e}}{\mathbf{s}} \right) \left[(1 - F^*(\mathbf{e} \mathbf{I} \mathbf{s}^{-1})) \right], \quad -\infty \leq \mathbf{e} \leq +\infty \quad (2.16)$$

onde: $\mathbf{s}^2 = \mathbf{s}_u^2 + \mathbf{s}_v^2$;

$$\mathbf{I} = \mathbf{s}_u / \mathbf{s}_v \text{ e}$$

$f^*(.)$ e $F^*(.)$ são as funções de densidade normal padronizada e normal cumulativa, respectivamente.

Segundo os mesmos autores, a referida densidade é assimétrica em torno de zero, com média e variância dadas por

$$E(\mathbf{e}) = E(u) = -\frac{\sqrt{2}}{\sqrt{p}} \mathbf{s}_u \quad (2.17)$$

$$V(\mathbf{e}) = V(u) + V(v) = \left(\frac{p-2}{p} \right) \mathbf{s}_u^2 + \mathbf{s}_v^2 \quad (2.18)$$

A parametrização seguida em (2.16) é considerada conveniente porque nela \mathbf{I} é interpretado como um indicador da variabilidade relativa das duas fontes de erro aleatório, as quais distinguem as firmas umas das outras. Quando $\mathbf{I}^2 \rightarrow 0$, implica em que $\mathbf{s}_v^2 \rightarrow \infty$ e/ou $\mathbf{s}_u^2 \rightarrow 0$, exprimindo o domínio do componente simétrico do erro na determinação de \mathbf{e} . Conseqüentemente, a equação (2.16) torna-se a densidade de uma variável aleatória $N(0, \mathbf{s}^2)$. Similarmente, quando $\mathbf{s}_v^2 \rightarrow 0$, o componente unicaudal do erro torna-se a fonte dominante da variação aleatória no modelo, e (2.13) toma a forma de uma seminormal negativa.

Nesta etapa de desenvolvimento, Aigner, Lovell e Schmidt colocam o problema de estimação, admitindo a disponibilidade de uma amostra aleatória com N observações e estabelecendo a seguinte função logarítmica de verossimilhança:

$$\begin{aligned} \ln J(y/\mathbf{b}, \mathbf{I}, \mathbf{s}^2) = & N \ln \frac{\sqrt{2}}{\sqrt{p}} + N \ln \mathbf{s}^{-1} \\ & + \sum_{i=1}^N \ln [1 - F^*(\mathbf{e}_i \mathbf{I} \mathbf{s}^{-1})] - \frac{1}{2\mathbf{s}^2} \sum_{i=1}^N \mathbf{e}_i^2 \end{aligned} \quad (2.19)$$

Considerando que para encontrar os valores dos estimadores \mathbf{b} , \mathbf{I} e \mathbf{s}^2 de (2.19), por MLE, já haviam algoritmos disponíveis, os pesquisadores não discutem o referido processo de cálculo. Observam que a função de verossimilhança é bem

comportada e apresentam as derivadas de segunda ordem usadas na obtenção dos erros-padrão assintóticos das estimativas.

De passagem, observam, também, que os coeficientes dos regressores podem ser estimados, sem viés e consistentemente, por OLS. E que os componentes de \mathbf{s}^2 podem ser extraídos dos resíduos de OLS, utilizando a equação (2.18) - em termos de \mathbf{s}_u^2 e \mathbf{s}_v^2 - e a sua relação com o terceiro momento dos resíduos de mínimos quadrados, que é

$$E[\mathbf{e} - E(\mathbf{e})]^3 = \frac{2\mathbf{s}_u^3}{\sqrt{2p}} \left(1 - \frac{4}{p}\right) \quad (2.20)$$

Com relação a distribuições alternativas para u_i , os autores em foco apresentam a função de densidade de $-u$, no caso de uma distribuição exponencial simples a um parâmetro, e ressaltam que todos os comentários e derivações desenvolvidos para a distribuição seminormal podem ser estendidos à distribuição exponencial.

Com o objetivo de descobrir alguma informação específica sobre o comportamento dos estimadores de máxima verossimilhança frente à pequenas amostras, eles constroem dois limitados experimentos tipo Monte Carlo com base em dados artificiais. Suas constatações, porém, não são conclusivas e eles reconhecem a necessidade da realização de outras pesquisas, inclusive para melhor avaliar a performance dos estimadores de momentos baseados nos resíduos de mínimos quadrados.

Finalmente, em um primeiro exemplo de aplicação empírica dos novos procedimentos metodológicos, Aigner, Lovell e Schmidt estimam por OLS, por programação linear (e quadrática) e por MLE funções de produção “média” e de fronteira, usando dados da indústria norte-americana de metais básicos referentes ao período 1957-58. Dentre outros resultados, indicam os valores das razões de t e de t assintóticos para os coeficientes de regressão, inclusive para \mathbf{s}_u^2 e \mathbf{s}_v^2 , tudo sob duas suposições para a distribuição de u (exponencial e seminormal). Com base nos valores maximizados dos logaritmos das funções de verossimilhança em cada caso, concluem que a distribuição seminormal proporciona um ajustamento marginalmente melhor. Na outra aplicação, usando dados da agricultura americana no período 1960-65, seguem quase os

mesmos procedimentos da aplicação anterior, porém sem utilizar a distribuição seminormal.

Nestas duas aplicações, é encontrada relativamente pouca representatividade do componente unicaudal com relação à S^2 , e os autores terminam o *paper* interrogando se os resultados permaneceriam para estabelecimentos individuais.

2.2 FRONTEIRA DE PRODUÇÃO AGREGADA

Aqui, emprega-se a denominação “fronteira de produção agregada” para designar a fronteira originária de funções de produção com múltiplos produtos e múltiplos insumos, em que os diferentes produtos são agregados em uma única variável dependente.

Sob a ótica da abordagem multi-produto, a nível de firma, a literatura sobre fronteiras estocásticas de produção tem considerado dois tipos de agregação: através de um índice (Lovell, 1995) e por intermédio dos coeficientes de correlação canônica (Arnold et al. 1996). A nível macroeconômico, não há referência na literatura sobre o enfoque multi-produto. Neste caso, os pesquisadores têm usado os preços dos produtos para calcular o valor adicionado ou o valor da produção, com a finalidade de utilizá-los como a variável dependente representativa do nível de produção agregado.

No Brasil, este tipo de fronteira agregada é o único encontrado na literatura. Assim procedem Pinheiro (1995) e Braga & Rossi (1986) em análises da eficiência técnica na indústria brasileira. No primeiro estudo, o autor observa que as estimativas do componente residual indicativo do nível de eficiência apresentam falta de assimetria, para todos os setores industriais analisados. Reconhecendo que a simetria dos resíduos unicaudais diminui consideravelmente o ganho de eficiência assintótica advindo da estimação por MLE, ele observa que este fenômeno aparece, também, em análises similares: na análise da indústria americana de metais realizada pelos idealizadores do modelo ALS, na investigação de dois dos cinco setores da indústria colombiana de transformação desenvolvida por Lee & Tyler (1978) e em 2/3 dos 136 setores industriais brasileiros analisados por Braga & Rossi (1986).

Se esta falta de significância estatística da distribuição do componente residual unicaudal é devida ao nível de agregação em que as análises são elaboradas ou não, é algo que ainda permanece para ser investigado. A este respeito, ainda, vale destacar que Schmidt & Lovell (1979) já haviam conjecturado que a evidência de eficiência técnica, ou de outro tipo, podia ser perdida no processo de agregação.

O enfoque de fronteira de produção agregada também é utilizado na análise da eficiência técnica da produção agropecuária conduzida por Nicolau (1983), sob o enfoque determinístico. Silveira et al. (1995) avaliam a eficiência técnica da clínica cirúrgica do Hospital Universitário da Universidade Federal de Santa Catarina, através de mínimos quadrados corrigidos, sob hipótese de uma fronteira também determinística. Já, na análise de 3 clínicas médicas do referido hospital, os mesmos pesquisadores valem-se de uma fronteira estocástica de custo, que é estimada por MLE (Silveira et al., 1997).

Assim, verifica-se que não há registro na literatura econométrica brasileira sobre o uso da fronteira estocástica de produção agregada, sob a ótica de múltiplos produtos conforme definida na presente pesquisa.

Segundo Lovell (1995), a agregação dos múltiplos produtos seguindo um índice é operacionalizada de acordo com o procedimento empregado por Gathon & Pestieau (1995). O referido estudo tem o objetivo de decompor a eficiência produtiva das empresas européias de transporte ferroviário, quanto à capacidade gerencial e ao grau de regulação. Para isto, os autores utilizam um modelo translog do tipo ALS, dado por

$$\ln \bar{y}(s, t) = \mathbf{a}_0 + \mathbf{a}_1 t + \frac{1}{2} \mathbf{a}_2 t^2 + \sum_m^4 \mathbf{a}_m \ln x_m(s, t) + \frac{1}{2} \sum_m^4 \mathbf{b}_m t \ln x_m(s, t) + \frac{1}{2} \sum_m^4 \sum_n^4 \mathbf{b}_{mn} \ln x_m(s, t) \ln x_n(s, t) + \mathbf{b}_0 \ln \bar{x}(s, t) + \mathbf{e}(s, t) \quad (2.21)$$

onde: \bar{y} é a soma dos dois produtos considerados;
 x_m são as variáveis representativas dos insumos ($m = 4$);
 \bar{x} é uma variável independente adicional dada por $\frac{y_1}{\bar{y}}$;
 t é uma variável temporal;
 (s, t) denotam a s -ésima cia. ferroviária e o t -ésimo ano,

respectivamente;

$\mathbf{a}'s$ e $\mathbf{b}'s$ são os coeficientes de regressão a estimar e
 $\mathbf{e}(s,t)$ é o termo residual composto pelo componente unicaudal
 $u(s,t)$ e pelo componente simétrico aleatório $v(s,t)$.

A equação (2.21) é estimada por MLE com o auxílio do LIMDEP, empregando um conjunto de dados referentes a 19 companhias ferroviárias em operação na Europa no período 1961-88. Sob a hipótese de seminormalidade, os $u(s,t)$ são estimados para cada ano e para cada empresa.

Apresentados como a média do período 1986-88, os índices de eficiência técnica variam de 0,947 (cia. NS/Holanda) a 0,742 (cia. NSB/Noruega). Para possibilitar a observação de mudanças temporais, são calculadas as diferenças entre os índices médios no período 1986-88 e no período inicial 1962-64. Também, são derivadas as taxas anuais de progresso técnico dos coeficientes associados com as variáveis de tendência e as taxas de variação da produtividade total dos fatores no período global de observação.

Na seqüência do estudo, Gathon e Pestieau, mediante pesquisa por questionário, constroem um índice representativo do grau de autonomia institucional. Com base neste índice, eles decompõem os escores de eficiência em eficiência gerencial e eficiência regulatória. Em resumo, suas conclusões indicam que os índices de eficiência e os *rankings* obtidos mudam bastante nas diferentes situações analisadas.

A outra maneira de agregar os múltiplos produtos deve-se a Arnold et al. (1996), que propõem uma nova combinação para o enfoque DEA/econométrico. Trabalhando com 5 produtos e 8 insumos referentes a 638 escolas públicas de nível secundário em atuação no Texas/EUA, eles obtêm resultados iniciais insatisfatórios. Neste desenvolvimento, estimam por OLS 5 funções do tipo Cobb-Douglas, em que cada produto Y é regredido sobre todos os insumos X_i ($i = 1,...,8$), ou seja,

$$\ln Y_i = \mathbf{b}_0 + \sum_{i=1}^8 \mathbf{b}_i \ln X_i + \mathbf{e}$$

.....

..... (2.22)

.....

$$\ln Y_s = b_0 + \sum_{i=1}^8 b_i \ln X_i + e$$

onde b_0 e b_i são os parâmetros a estimar.

Este procedimento é realizado, principalmente, para verificar a significância estatística dos coeficientes de regressão e a direção de seus correspondentes sinais. O resultado da referida verificação aponta que apenas o coeficiente de um dos insumos mostra-se estatisticamente significativo e com o mesmo sinal em todas as regressões.

Mesmo assim, os pesquisadores reestimam as 5 funções, porém num contexto de fronteira estocástica nos moldes do modelo ALS, supondo que $e = v - t$ com v irrestrito quanto ao sinal e que $t \geq 0$ é distribuído seminormalmente. Seguindo este modelo eles definem a função logarítmica de máxima verossimilhança para as funções acima, que são maximizadas por MLE para estimação dos parâmetros. As estimativas de eficiência técnica para cada escola são, então, estimadas através de $\exp(-\hat{t})$ de acordo com o método de Jondrow et al. (1982) detalhado na seção 3.1.3.5.

Assim, a eficiência média para todas as escolas é estimada em 0,97, com índice individual mínimo igual a 0,88 e máximo a 0,99. Através deste *rank* não é possível distinguir com clareza as escolas eficientes e as ineficientes. Estes resultados decorrem dos problemas de significância e de sinal ocorridos na estimação por OLS, cujos coeficientes de regressão obtidos, à exceção do intercepto, são os mesmos do modelo de fronteira estocástica.

Para verificar se a origem destas deficiências é o uso das regressões com uma variável dependente de cada vez, os pesquisadores passam a conceitualizar a educação como um processo de produção em que todos os produtos são produzidos conjuntamente em um único processo. Segundo eles, na teoria econômica formal, uma função de produção conjunta é representada da forma

$$F(Y_1, Y_2, \dots, Y_s ; X_1, X_2, \dots, X_m) = 0 \quad (2.23)$$

quando a UTD produz s produtos conjuntos, Y_1, Y_2, \dots, Y_s , utilizando m insumos, X_1, X_2, \dots, X_m .

Para modelar estes s produtos conjuntos em uma única função de produção e estimar os coeficientes, eles reescrevem (2.23) como

$$f(Y_1, Y_2, \dots, Y_s) = g(X_1, X_2, \dots, X_m) e^e \quad (2.24)$$

cujas especificação na forma Cobb-Douglas estendida é

$$Y_1^{a_1} Y_2^{a_2} \dots Y_s^{a_s} = b_0 X_1^{b_1} X_2^{b_2} \dots X_m^{b_m} e^e \quad (2.25)$$

onde os a 's e os b 's representam os parâmetros a serem estimados.

Esta função é estimada pelo método CCA, no qual r representa a correlação canônica entre os vetores Y e X , ou seja,

$$r = r_{vu} \quad \text{para} \quad v = a'Y \quad \text{e} \quad u = b'X \quad (2.26)$$

onde: Y e X são padronizados de maneira que a soma das variâncias de cada um é a unidade;

a' e b' são vetores-linha cujos componentes, chamados de coeficientes canônicos, são determinados quando r é maximizado e

v e u são os resultados de produtos escalares, sendo denominados de variáveis canônicas para Y e X , respectivamente.

Desta forma, a estimação de (2.25) é obtida por

$$\hat{a} = a \quad \text{e} \quad \hat{b} = br \quad (2.27)$$

onde r é a mais alta raiz das equações características formadas pelas observações.

Na sequência do estudo, Arnold e sua equipe utilizam os \hat{a} 's oriundos desta estimação para definir um produto agregado e expressar (2.25) como

$$\ln Y = \ln b_0 + \sum_{i=1}^m b_i \ln X_i + v - t \quad (2.28)$$

onde $\ln Y = \sum_{r=1}^s \hat{a}_r \ln Y_r$ é definida como uma nova variável dependente.

Esta fronteira estocástica é, então, estimada por MLE (fronteira estocástica) para todas as escolas da amostra e, por OLS, para cada um dos dois subconjuntos de escolas previamente identificadas como ineficientes e eficientes, segundo a análise DEA. Resultados satisfatórios são, finalmente, colhidos, através do uso conjunto dos métodos DEA e econométrico. A amostra inteira é submetida a uma única regressão com variáveis

dummy para distinguir entre escolas DEA-eficientes e escolas DEA-ineficientes. Os coeficientes dos insumos são positivos para as escolas eficientes e negativos e estatisticamente significantes para as escolas ineficientes. Estes resultados são considerados consistentes com a teoria econômica e de utilidade para a política educacional.

2.3 FRONTEIRA DE CUSTO

Schmidt & Lovell (1979) fazem uma extensão do modelo ALS de função de produção de fronteira estocástica, explorando a dualidade entre as funções de produção e de custo, sob as suposições de exata e de inexata minimização de custos. Eles desenvolvem uma metodologia para medir os diferentes tipos de ineficiência e os custos de ineficiência associados, ilustrando as referidas técnicas através de uma aplicação empírica. Como a presente pesquisa está delimitada ao âmbito exclusivo da eficiência técnica, faz-se a revisão do referido *paper* somente no que diz respeito a este enfoque.

Schmidt e Lovell consideram que a firma procura minimizar o custo de produzir uma desejada quota de produção, sujeita a restrições de uma fronteira de produção estocástica, e que ela pode ser tecnicamente ineficiente e operar abaixo dessa fronteira - mas deve ser alocativamente eficiente, ou seja, operar no seu caminho de expansão a um custo mínimo¹. Assim, eles caracterizam a tecnologia de produção da firma pela função de produção dada por

$$y = a \prod_{i=1}^n x_i^{a_i} e^e \quad (2.29)$$

onde: y é o produto da firma;

x_i são os insumos para o processo de produção;

$e = v - u$ é um distúrbio aleatório (como no modelo ALS) e

a e a_i são parâmetros a estimar.

A partir da forma logaritmada

¹ Aqui parece haver uma inconsistência. Se a firma estiver operando com eficiência alocativa ela, obrigatoriamente, deve estar operando com eficiência técnica.

$$\ln y = A + \sum_{i=1}^n \mathbf{a}_i \ln x_i + (v - u) \quad (2.30)$$

onde $A = \ln a$, eles notam que $\ln y$ é limitada, superiormente, pela fronteira de produção estocástica

$$A + \sum_{i=1}^n \mathbf{a}_i \ln x_i + v \quad (2.31)$$

em que a eficiência técnica com relação à fronteira é dada na proporção u .

Como cada firma é alocativamente eficiente, ela não comete erro na seleção das proporções de fatores que minimizam o custo, as quais são dadas pela solução de

$$\ln x_i - \ln x_i = B_i, \quad i = 2, \dots, n, \quad (2.32)$$

onde: $B_i = \ln(p_i \mathbf{a}_i / p_1 \mathbf{a}_1)$ e

p_1, p_2, \dots, p_n são os preços dos insumos.

A partir daqui, Schmidt e Lovell derivam as equações da demanda por fator

$$\ln x_i = \ln k_i + \frac{1}{r} \ln y + \ln \left[\prod_{j=1}^n p_j^{a_j/r} / p_i \right] - \frac{1}{r} (v - u), \quad i = 1, \dots, n \quad (2.33)$$

e chegam a uma função de custo da forma

$$\ln C = K + \frac{1}{r} \ln y + \sum_{i=1}^n \frac{\mathbf{a}_i}{r} \ln p_i - \frac{1}{r} (v - u) \quad (2.34)$$

onde: $K = \ln \left[\sum_{i=1}^n k_i \right] = \ln r - \frac{1}{r} A - \frac{1}{r} \ln \left[\prod_{i=1}^n \mathbf{a}_i^{a_i} \right]$ e

$$K + \frac{1}{r} \ln y + \sum_{i=1}^n \frac{\mathbf{a}_i}{r} \ln p_i - \frac{1}{r} v$$

é a fronteira estocástica de

custo que limita, inferiormente,

$\ln C$.

Na equação (2.34), $r = \sum_{i=1}^n \mathbf{a}_i$ mede os retornos de escala, e $(1/r) u$ mede o

montante de custo extra para o produtor, em decorrência dele ter produzido abaixo da fronteira de produção. Ou seja, $(1/r) u$, é a estimativa da ineficiência técnica, supondo que o produtor seja alocativamente eficiente. Assim, uma firma pode estar acima de seu

custo de fronteira, somente se estiver abaixo de sua fronteira de produção. Nestas condições, o custo de tal ineficiência técnica varia inversamente com os retornos de escala.

Schmidt e Lovell procedem à análise empírica de empresas geradoras de energia elétrica, em que o nível de produto é exógeno à firma, a qual maximiza lucros simplesmente pela minimização do custo de produzir um determinado volume de produção. Supondo que os preços dos insumos são também fixados para a firma, o modelo fica com a quantidade de produção e os preços dos insumos como variáveis exógenas, e as quantidades de insumos e o custo, como endógenas. Como todas as variáveis do lado direito da função de custo são exógenas, a equação pode ser estimada sem preocupação com complicações devidas a simultaneidade e, sendo linearmente homogênea nos preços dos insumos, pode ser reescrita como

$$\ln(C/p_n) = K + \frac{1}{r} \ln y + \sum_{i=1}^{n-1} \frac{a_i}{r} \ln(p_i/p_n) - \frac{1}{r} (v - u) \quad (2.35)$$

Fazendo a suposição de que a distribuição de u é seminormal, eles mostram que a presença deste componente no resíduo aleatório é manuseada analogamente ao modelo ALS, havendo apenas a necessidade de mudar alguns sinais.

Para estimar (2.35), Schmidt e Lovell utilizam uma amostra de 111 novas empresas privadas geradoras de energia elétrica, construídas nos Estados Unidos entre 1947 e 1965, cujos dados além do custo e da produção de eletricidade, incluem observações sobre o uso dos 3 insumos principais. Em um dos métodos de estimação adotados, eles obtêm por OLS estimativas consistentes dos parâmetros $1/r$, a_1/r , ..., a_{n-1}/r , de onde extraem r , a_1 , a_2 , ..., a_n . A seguir, calculam o segundo e o terceiro momentos dos resíduos de OLS, \hat{m}_2 e \hat{m}_3 , que são estimativas consistentes, respectivamente, de

$$m_2 = \left[s_v^2 + \frac{p-2}{p} s_u^2 \right] / r^2 \quad \text{e de} \quad m_3 = \left\{ \sqrt{\frac{2}{p}} \left[\frac{4-p}{p} \right] s_u^3 \right\} / r^3 \quad (2.36)$$

de onde deduzem

$$s_u^2 = \left[\sqrt{\frac{p}{2}} \left(\frac{p}{4-p} \right) m_3 \right]^{2/3} r^2 \quad \text{e} \quad s_v^2 = r^2 m_2 - \left(\frac{p-2}{p} \right) s_u^2 \quad (2.37)$$

Então, usando as estimativas consistentes $\hat{\mathbf{m}}_2$ e $\hat{\mathbf{m}}_3$, obtidas dos resíduos de OLS, no lugar de \mathbf{m}_2 e \mathbf{m}_3 em (2.37), eles conseguem estimativas consistentes de $\hat{\mathbf{s}}_u^2$ e de $\hat{\mathbf{s}}_v^2$.

Como último procedimento de estimação, valem-se da equação

$$E(\hat{K}) = \text{plim } \hat{K} = K + \frac{1}{r} E(u) = K + \frac{1}{r} \sqrt{\frac{2}{\mathbf{p}}} \mathbf{s}_u \quad (2.38)$$

e das estimativas consistentes \hat{r} de r e $\hat{\mathbf{s}}_u$ de \mathbf{s}_u para explicitar a estimativa consistente do coeficiente de interseção corrigido \tilde{K} , como sendo

$$\tilde{K} = \hat{K} - \frac{1}{\hat{r}} \sqrt{\frac{2}{\mathbf{p}}} \hat{\mathbf{s}}_u \quad (2.39)$$

onde \hat{K} é a estimativa de OLS de K .

Assim, são obtidas estimativas consistentes para todos os parâmetros do modelo, através do que os dois autores designam estimadores OLS/momentos, que mais tarde passam a integrar a literatura com o rótulo de estimadores COLS.

Como segunda possibilidade de estimar a fronteira de custo, os referidos autores empregam MLE, com uma função de verossimilhança muito similar à definida no modelo ALS.

Maximizando a referida função, Schmidt e Lovell obtêm as estimativas de máxima verossimilhança de todos os parâmetros da função de custo e das distribuições de u e de v , ou seja, K , $1/r$, \mathbf{a}_1/r , ..., \mathbf{a}_{n-1}/r , $\mathbf{I} = \mathbf{s}_u/\mathbf{s}_v$ e $\mathbf{s}_\bullet^2 = (1/r^2)[\mathbf{s}_v^2 + \mathbf{s}_u^2]$.

Com relação aos resultados alcançados, os pesquisadores observam que o método OLS/momentos dá uma estimativa negativa de \mathbf{s}_v^2 e que os demais resultados são completamente razoáveis. Também, notam que a média do resíduo unicaudal na função de custo, $(1/r)u$, é 0,1260, significando que o custo está, em média, 12,6% acima do seu nível na fronteira. A esta última medida, chamam “custo da ineficiência técnica”.

Sobre o emprego de OLS/momentos ou de MLE como métodos de estimação, eles elegem como principal vantagem do primeiro uma maior facilidade de computação e, como sua principal desvantagem, o fato de produzir estimativas menos eficientes do que MLE. Além disso, não existe garantia de que as estimativas de \mathbf{s}_u^2 e de \mathbf{s}_v^2 , por

OLS/momentos, sejam não-negativas. Se o terceiro momento dos resíduos de OLS, $\hat{\mathbf{m}}_3$, é negativo, $\hat{\mathbf{S}}_u$ também o é. Ainda que $\hat{\mathbf{S}}_u^2$ seja positiva, isto é problemático. Schmidt e Lovell alertam, também, para a possibilidade de encontrar $\hat{\mathbf{m}}_2$ e $\hat{\mathbf{m}}_3$, tais que $\hat{\mathbf{S}}_v^2 < 0$. Esta situação pode ser prevenida com o uso de MLE e, quando se está certo da especificação adotada, isto pode ser uma vantagem adicional deste método. Por outro lado, pode ser argumentado, em contrapartida, que é uma vantagem do método OLS/momentos o fato de se conhecer previamente a ocorrência de um $\hat{\mathbf{m}}_3$ negativo e, imeditamente, reavaliar o erro de especificação.

Seguindo a linha metodológica de Schmidt & Lovell (1979), dez anos após, Kumbhakar (1989) estende aqueles resultados em duas direções principais. Em uma, estima a ineficiência técnica para cada firma e, em outra, incorpora à análise, múltiplos produtos e fatores de produção exógenos.

Para estimar apenas a ineficiência técnica, Kumbhakar trabalha com uma bem comportada tecnologia de produção, que é especificada como

$$f(\mathbf{y}) = g(X, Z) \exp(\mathbf{t} + v) \quad (2.40)$$

onde: \mathbf{Y} é o vetor de produtos;
 X é o vetor de insumos endógenos;
 Z é o vetor de insumos exógenos;
 v e \mathbf{t} são variáveis aleatórias (como no modelo ALS);
 $f(\mathbf{y})$ é a função de produção e
 $g(X, Z)$ é a função dos insumos.

Supondo f e g do tipo Cobb-Douglas e introduzindo subscritos para o tempo e para as firmas, ele reescreve (2.40), na forma logarítmica, como

$$\sum_{i=1}^s d_i \ln \Psi_{ift} = a_0 + \sum_{j=1}^n a_j \ln X_{ift} + \sum_{k=1}^m g_k \ln Z_{kft} + \mathbf{t}_{ft} + v_{ft} \quad (2.41)$$

onde o subscrito $f = 1, \dots, F$, indexa as firmas, e $t = 1, \dots, T$, indexa o tempo.

Apesar desta especificação da tecnologia de produção não satisfazer as condições usuais de convexidade² quanto à maximização de lucros, ela foi utilizada porque os produtos são exógenos, no caso das ferrovias estudadas. Como prestadoras de serviços públicos, as empresas ferroviárias oferecem transporte por demanda da clientela, sob tarifas e condições bem determinadas, objetivando, a curto prazo, operar com o mínimo custo.

Levando em conta esta concepção e admitindo exogeneidade nos preços dos insumos e nas quantidades dos produtos, o autor define as condições de primeira ordem para minimização de custos, frente à escolha de diferentes níveis de insumos endógenos, como sendo

$$\ln X_{ift} - \ln X_{1ft} + \ln W_{ift} - \ln W_{1ft} = \ln(\mathbf{a}_j / \mathbf{a}_1), \quad j = 2, 3, \dots, n \quad (2.42)$$

onde $W_j (j = 1, 2, \dots, n)$ são os preços dos insumos endógenos X_j .

Como esta equação é completamente satisfeita para todo j, f e t , as firmas são alocativamente eficientes. Assim, pode-se substituir (2.42) em (2.41), para fins de estimação, e obter

$$\sum_i d_i \ln y_{ift} = \mathbf{a}_0 + \mathbf{a}_1 X_{ft}^* + \sum_k \mathbf{g}_k \ln Z_{kft} + \mathbf{t}_{ft} + v_{ft} \quad (2.43)$$

onde $X_{ft}^* = \ln X_{1ft} + \sum_j \left(X_{ift} W_{ift} / X_{1ft} W_{1ft} \right) \ln X_{ift}$.

Então, sob a suposição de exata minimização de custos, Kumbhakar indica dois procedimentos alternativos para estimar os parâmetros de interesse. Em um deles, aplica OLS à equação

$$X_{ft}^* = \left(\sum_i d_i \ln y_{ift} - \mathbf{a}_0 - \sum_k \mathbf{g}_k \ln Z_{kft} \right) / \mathbf{a}_1 - (\mathbf{t}_{ft} + v_{ft}) / \mathbf{a}_1 \quad (2.44)$$

que é a forma estimável de (2.43) no caso de X^* ser endógeno, obtendo estimativas consistentes de \mathbf{a}_1 , d_i e \mathbf{g}_k . Os parâmetros remanescentes $(\mathbf{a}_2, \dots, \mathbf{a}_n)$ são estimados consistentemente por (2.42), dado que \mathbf{a}_1 é previamente estimado por (2.44).

² Segundo o mesmo autor, este problema pode ser evitado definindo uma função com elasticidade de substituição constante para os insumos e uma função com elasticidade de transformação constante para os produtos.

O outro procedimento de estimação recomendado é o emprego da seguinte forma modificada de (2.34), para acomodar múltiplos produtos e insumos exógenos:

$$\ln C_{jft} = A + \frac{1}{r} \sum_i d_i \ln y_{ifft} + \sum_i \frac{a_i}{r} \ln W_{ifft} - \sum_k \frac{g_k}{r} \ln Z_{kft} - \frac{1}{r} (v_{jft} + t_{jft}) \quad (2.45)$$

onde: $C = \sum_i W_i X_i$;

$$r = \sum_i a_i \text{ e}$$

A é uma bem definida função de a_0, a_1, \dots, a_n .

Como todas as variáveis do lado direito desta equação são exógenas, seus parâmetros podem ser estimados consistentemente por OLS. E, a partir do segundo e do terceiro momento dos resíduos da regressão, são encontradas as estimativas de s_t^2 e de s_v^2 .

Apesar de ter enfatizado estes dois caminhos para o trabalho de estimação, Kumbhakar prefere utilizar estimadores MLE. Com esta finalidade, ele deriva e maximiza a função logarítmica de verossimilhança de (2.43), mencionando que um procedimento similar pode ser adotado para a função de custo (2.45). Obtem, então, as estimativas de máxima verossimilhança dos parâmetros da função de produção, bem como de s_t^2 e s_v^2 (que são consistentes e eficientes). Dado que a_1 fica estimado, a estimação de a_2, \dots, a_n é obtida de (2.42).

Para estimar a magnitude da ineficiência técnica de cada firma, o autor deriva a função densidade de probabilidade de t dado e , $g(t / e)$, considerando $t|e$ como uma $N(m, s^2)$ truncada em zero. A estimativa de t para cada firma, de acordo com a solução de Jondrow et al. (1982) apresentada na seção 3.1.3.5, é definida com base na moda. Uma vez efetuada a estimação de t , as estimativas do custo da ineficiência técnica para cada UTD são obtidas por meio de $-\frac{1}{r} t$.

Utilizando a metodologia descrita até aqui, e um painel de dados referente ao período 1951-75, Kumbhakar estima a ineficiência técnica de 13 empresas ferroviárias norte-americanas. Seu vetor de produtos é integrado por duas variáveis de produção

(número de passageiros e toneladas de carga transportados) e por duas outras indicativas da qualidade dos serviços prestados (distância média das viagens para transporte de cargas e de passageiros). No vetor de insumos, usa variáveis referentes ao emprego dos fatores capital, trabalho e combustível.

Na fase de interpretação de resultados, o pesquisador explica que o índice de eficiência técnica multiplicado por 100 indica a porcentagem pela qual o nível de produção da empresa pode ser aumentado, desde que ela deixe de ser tecnicamente ineficiente. Como os produtos são tratados como exógenos, por causa do enfoque de minimização de custos, ele argumenta que é mais apropriado observar o efeito da ineficiência técnica sobre a demanda por insumos e interpretar a ineficiência técnica sob o ângulo do seu custo. Com base no modelo de Schmidt & Lovell (1979), Kumbhakar coloca que o impacto da ineficiência técnica sobre a demanda de insumos, bem como sobre o custo, é a elevação destas variáveis em $-\frac{1}{r}t$ 100%. Este é o índice do custo da ineficiência técnica, cujo entendimento o pesquisador ilustra usando um valor igual a 0,29 auferido por uma determinada companhia. Neste caso, o índice indica que a demanda pelos insumos utilizados está aumentada em 29%, devido à ineficiência técnica com que a referida empresa está operando. Em outras palavras, equivale a dizer que o referido custo poderia ficar 29% mais baixo, desde que a empresa resolvesse o seu problema de ineficiência.

No tocante à tarefa de verificação, Kumbhakar testa a suposição de exata minimização de custos através da verificação dos valores da razão $(W_j X_j / W_L L)$ para os insumos capital e combustível. Como para ambos os insumos a diferença entre o mínimo e o máximo valor é de, aproximadamente, cinco vezes, a hipótese de exata minimização de custos é rejeitada. Segundo o autor, existem boas razões para que uma indústria, regulada como a ferroviária, opere com ineficiência alocativa. Ele, também, fornece os valores dos erros-padrão assintóticos dos coeficientes de regressão estimados, porém não explicita nada sobre a significância estatística dos mesmos.

Quando a hipótese de minimização de custos não é verdadeira, as estimativas dos parâmetros da função de produção obtidos através da função de custo (2.45) são consistentes, mas não eficientes. Conseqüentemente, não se pode recorrer à propriedade

de invariância dos estimadores MLE, para garantir que eles sejam de fato os estimadores MLE oriundos da maximização de (2.43). Neste caso, Kumbhakar aponta como estratégia metodológica a estimação por MLE de um sistema de equações simultâneas, integrado pela função de produção (2.41) e pelas condições de primeira ordem para a minimização de custos, dadas por

$$\ln X_{ift} - \ln X_{1ft} + \ln W_{ift} - \ln W_{1ft} = \ln(\mathbf{a}_j / \mathbf{a}_1) + \mathbf{x}_{ift}, \quad j = 2, \dots, n \quad (2.46)$$

onde \mathbf{x}_{ift} é a ineficiência alocativa que representa o montante pelo qual o par de insumos $(1, j)$ difere do par de mínimo custo da firma f no período t .

Já a alternativa de estimação via função de custo, é considerada pelo pesquisador como mais complicada e não muito recomendável. Entretanto, ele reconhece que sempre é possível estimar o custo da ineficiência técnica e da ineficiência alocativa, quando os parâmetros são estimados com base em uma função de custo. Kumbhakar não vai além destas colocações, porém é importante procurar explorar esta questão um pouco mais.

A maior dificuldade na análise de eficiência técnica, utilizando funções de custo alheias à hipótese de minimização de custos, é a decomposição dos desvios da fronteira em medidas de eficiência técnica e de eficiência alocativa.

Empregando um sistema de equações constituído pela função de custo e pelas equações associadas da demanda por insumos, Kopp & Diewert (1982) haviam desenvolvido um método pelo qual a referida decomposição é realizada. Este método, porém, não especifica a exata relação entre o componente residual unicaudal da função de custo e o componente residual bicaudal das equações de participação dos insumos. Este problema, conhecido na literatura como “*Greene problem*”, ainda não está completamente solucionado. O último avanço nesta direção é a contribuição de Mensah (1994), formulada em um trabalho destinado a propor e apresentar uma simplificação do método original da decomposição.

Diante desta realidade, é compreensível que quase todos os estudos econométricos através de funções de custo de fronteira, tratem exclusivamente da estimação da eficiência produtiva, sem decompô-la. A medida conseguida com este

procedimento é designada, unanimemente, índice de eficiência do custo (da eficiência técnica e da eficiência alocativa, conjuntamente).

2.4 FRONTEIRA DE PRODUÇÃO INVERSA

Gathon & Perelman (1992) adotam enfoque baseado na estimação de uma função de necessidades de fator, também conhecida como função de produção inversa (Diewert, 1974). Estimam os indicadores de eficiência técnica de 19 companhias ferroviárias da Europa no período 1961-1988, através de um painel de dados. Na função utilizada, as unidades de trabalho são expressadas como a variável dependente, requerendo uma suposição de complementariedade (proporções fixadas) entre todos os insumos principais, para ser uma representação válida do processo produtivo. Por isso, ao dispor de uma amostra em que as despesas com pessoal perfazem cerca de 90% do custo variável, os pesquisadores aceitam que as possibilidades de substituição entre os insumos trabalho e energia são altamente limitadas.

Reconhecendo que embora esta abordagem não seja o caminho usual para modelar atividades de produção, eles argumentam que ela parece um campo promissor para empresas que produzem múltiplos produtos e operam em condições de excessiva regulamentação. Também, ressaltam que ela pode ser preferível à função de produção usual, devido aos problemas para obter formas funcionais com múltiplos produtos.

A função empregada por eles é da forma

$$\begin{aligned} y_{it} &= \mathbf{a}_0 + X'_{it} \mathbf{b} + r_t \mathbf{t} + Z'_i \mathbf{g} + \mathbf{e}_{it} \\ \mathbf{e}_{it} &= u_i + v_{it} \end{aligned} \quad (2.47)$$

onde: y_{it} , X'_{it} e Z'_i são variáveis expressas em logaritmos na turais;
 $i = 1, \dots, I$ e $t = 1, \dots, T$ indexam firmas e períodos, respectivamente;
 y_{it} é o insumo endógeno;
 X_{it} é o vetor de variáveis exógenas;
 r_t e Z_i são as variáveis temporais e de *cross-section*, respectivamente;

\mathbf{a}_0 , \mathbf{b} , \mathbf{t} e \mathbf{g} são os parâmetros a estimar e
 \mathbf{e}_{it} é o termo residual composto por u_i , que mede os efeitos individuais invariáveis no tempo e v_{it} , que, por hipótese, é normalmente distribuído e não-correlacionado com u_i e os regressores.

A especificação (2.47), desde que $u_i \geq 0$, é um caso especial do modelo ALS, em que a única diferença reside no painel de dados, que oferece um caminho natural para distinguir o indicador de eficiência e o ruído estatístico.

Para possibilitar a aplicação do método de painel, (2.47) é transformada em

$$y_{it} = X'_{it}\mathbf{b} + r_i\mathbf{t} + Z'_i\mathbf{g} + \mathbf{a}_i + v_{it} \quad (2.48)$$

onde $\mathbf{a}_i = \mathbf{a}_0 + u_i$.

Esta equação é estimada por quatro diferentes métodos: efeitos fixos (WITHIN), efeitos aleatórios (GLS), mínimos quadrados generalizados com variável instrumental (GLS-IV) e MLE. A estimação de (2.48) por MLE é realizada sob a suposição de que a eficiência técnica u_i é seminormalmente distribuída, seguindo o desenvolvimento de Battese e Coelli (1988) detalhado na seção 3.1.3.5. A verificação da significância estatística dos coeficientes estimados através dos métodos WITHIN, GLS, GLS-IV e MLE é feita através da estatística t . Os índices de eficiência obtidos segundo cada um dos métodos são comparados por intermédio do coeficiente de correlação ordinal de Spearman, cuja significância é testada por χ^2 .

Dentre as conclusões emitidas, Gathon e Perelman destacam os altos graus de correlação entre os índices de eficiência técnica, a comprovação da influência de economias de escala e a constatação de que a função utilizada é uma maneira simples e conveniente para modelar a atividade produtiva de companhias ferroviárias, que são altamente reguladas e não oferecem grandes possibilidades de substituição entre os insumos.

Em uma outra contribuição, Kumbhakar & Hjalmarsson (1995) empregam uma função de produção inversa para analisar a eficiência de seguradoras suecas no período

1974-84, com base em um grande painel de dados com 380 observações em cada ano. Com a justificativa de que o custo da mão-de-obra perfazia cerca de 80% do custo total dos escritórios, e que estes produziam múltiplos produtos (serviços), eles concentram o estudo na eficiência do uso do fator trabalho.

Com o objetivo de impor o mínimo de restrições à tecnologia, adotam uma formulação translog do tipo

$$\ln L = \mathbf{b}_0 + \sum_i \mathbf{b}_i \ln Y_i + \frac{1}{2} \sum_i \sum_j \mathbf{b}_{ij} \ln Y_i Y_j + \mathbf{t} + v \quad (2.49)$$

onde: L é o insumo trabalho;
 Y_1, \dots, Y_n são os n serviços produzidos usando mão-de-obra e
 $\mathbf{b}_0, \mathbf{b}_i$ e \mathbf{b}_{ij} são os parâmetros a estimar.

Para atender a metodologia de painel de dados, a estrutura dos resíduos de (2.49) é especificada como

$$\mathbf{e}_{ft} \equiv \mathbf{t}_{ft} + v_{ft} = \mathbf{m}_f + \mathbf{I}_t + u_{ft} + v_{ft} \quad (2.50)$$

onde: f e t são os indexadores das firmas e do tempo ($f = 1, \dots, F$, $t = 1, \dots, T$),
respectivamente;
 \mathbf{t}_{ft} é o componente do termo \mathbf{e}_{ft} indicativo da ineficiência;
 \mathbf{m}_f indica a ineficiência técnica específica de cada firma;
 \mathbf{I}_t é o componente de variação no tempo e
 u_{ft} é o componente residual unicaudal indicativo da ineficiência técnica.

Como fazem parte da ineficiência, todos estes componentes são não-negativos. Por sua vez, o componente v_{ft} tem papel similar a v_i no modelo ALS.

Na estimação de (2.49), Kumbhakar e Hjalmarsson utilizam dados do fator trabalho e de 4 tipos de produtos (serviços) e preferem a estimação por múltiplos estágios, via OLS e otimização de funções de probabilidade. Mesmo reconhecendo que o método MLE é direto, deixam de usá-lo devido à existência de muitos parâmetros e de um grande número de unidades de *cross-section*. Para estimar u_{ft} , seguem a abordagem de Jondrow et al. (1982) - ver seção 3.1.3.5 - e calculam os índices de eficiência no uso do fator trabalho, para cada firma e para cada ano, através de

$$\hat{TE}_{jt} = \exp(-\hat{t}_{jt}) \quad (2.51)$$

onde $\hat{t}_{jt} = \hat{I}_t + \hat{m}_j + \hat{u}_{jt}$.

Na investigação de alguns requisitos estatísticos e econométricos, os pesquisadores constataam que todos os b'_{it} são positivos e estatisticamente significantes. A hipótese de normalidade dos resíduos é rejeitada pelo teste Kolmogorov-Smirnov. A estatística de qui-quadrado é empregada para avaliar a qualidade do ajustamento de u a uma distribuição seminormal. Esta suposição não é aceita ao nível de 5%.

Os dois autores também realizam análises de sensibilidade para avaliar o impacto decorrente de mudanças na distribuição de u . Isto é feito mediante a maximização das funções logarítmicas de verossimilhança, em cada caso. A conclusão alcançada aponta que as estimativas de eficiência são bastante robustas quanto à mudança na especificação de u , de seminormal para exponencial. Em outra etapa da verificação, são testadas, via teste de Chow, as diferenças encontradas nos b'_{it} estimados no modelo com m_j e I_t , em relação aos mesmos coeficientes estimados no modelo sem os referidos componentes. Como resultado, o modelo sem m_j e I_t é rejeitado ao nível de 1%.

Dentre os principais resultados empíricos, sobressaem a comprovação da existência de retornos crescentes de escala no período 1974-84, a evidência de um suave decréscimo na produtividade da mão-de-obra ao longo do tempo e a variação da eficiência média no uso do fator trabalho, entre 77,42% e 82,85%.

2.5 FRONTEIRA DE DISTÂNCIA-PRODUTO

Lovell et al. (1990) empregam uma função de distância-produto e constroem um índice sobre a eficiência da população australiana em transformar recursos materiais e financeiros em bem-estar.

Neste trabalho, as variáveis de interesse são representadas por um vetor de recursos, $x = (x_1, \dots, x_N) \in \mathfrak{R}_{++}^N$, por um vetor dos níveis de satisfação, $u = (u_1, \dots, u_M) \in \mathfrak{R}_{++}^M$ e por um vetor englobando o conjunto dos recursos e dos níveis de satisfação individuais, (x^i, u^i) , $i = 1, \dots, I$.

A função de distância -produto é definida como

$$D_0(x, u) = \min \{q: u/q \in P(x)\} \quad (2.52)$$

onde: $P(x)$ é o conjunto de todos os vetores de níveis de satisfação que podem ser realizados com o vetor de recursos x ;

$D_0(x, u)$ é uma função não-decrescente, positiva e linearmente homogênea e convexa em u , e decrescente em x e

$D_0(x, u) \leq 1$ se $u \in P(x)$, tendo igualdade mantida se $u \in \text{Isoq } P(x) = \{u: u \in P(x), \text{ } \mathbf{1}u \notin P(x), \text{ } \mathbf{1} > 1\}$.

O indicador de eficiência utilizado é o índice de produtividade de Malmquist

$$P(x^k, x^l, u^k, u^l) = \frac{D_0(x^k, u^k)}{D_0(x^l, u^l)} \quad (2.53)$$

onde (x^k, u^k) e (x^l, u^l) são dois diferentes vetores de recursos e de níveis de satisfação.

Neste caso, o conjunto de referência que limita os dados por cima é a isoquanta

$$\text{Isoq } P(x^i) = \left\{ (u^i/q_i): D_0(x^i, u^i/q_i) = 1, \quad i = 1, \dots, I \right\} \quad (2.54)$$

Para fins de estimação, os autores valem-se das propriedades de homogeneidade e de fechamento, e assumem uma forma translog do tipo

$$\begin{aligned}
\ln(u_M^{-1}) = & \mathbf{a}_0 + \sum_{i=1}^N \mathbf{a}_i \ln x_i + \sum_{i=1}^{M-1} \mathbf{b}_i \ln \hat{u}_i \\
& + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \mathbf{a}_{ij} \ln x_i \ln x_j + \frac{1}{2} \sum_{i=1}^{M-1} \sum_{j=1}^{M-1} \mathbf{b}_{ij} \ln \hat{u}_i \ln \hat{u}_j \\
& + \sum_{i=1}^N \sum_{j=1}^{M-1} \mathbf{g}_{ij} \ln x_i \ln \hat{u}_j + \mathbf{e}
\end{aligned} \tag{2.55}$$

onde $\hat{u}_i = (u_i / u_M)$ e $\hat{u}_j = (u_j / u_M)$, $i, j = 1, \dots, M-1$, $i \neq j$.

Esta equação é estimada com uma amostra de dados levantados pela pesquisa do padrão de vida dos australianos realizada em 1987. É empregado o método dos mínimos quadrados ordinários modificado (*modified ordinary least squares* - MOLS) para garantir que todos os vetores individuais de recursos e de níveis de satisfação fiquem sobre ou abaixo da isoquanta de fronteira estimada. Dos resíduos modificados são extraídos os valores da função de distância-produto para cada indivíduo, por meio da transformação

$$D_0(x^k, u^k) = \exp \{ (\text{resíduo negativo máximo}) - (\text{resíduo})^k \} \tag{2.56}$$

onde $k = 1, \dots, I$.

Estes valores medem a habilidade com que os indivíduos convertem seus recursos em bem-estar. Dividindo cada um dos valores $D_0(x^k, u^k)$ pelo máximo valor dentre todos eles que, por construção, é a unidade, são gerados os respectivos índices de eficiência de transformação para cada indivíduo da amostra, ou seja, $P(x^k, x^l, u^k, u^l)$.

Como conclusão, os autores destacam que o índice medido, ainda que altamente correlacionado com o índice do nível de satisfação, como de fato deve ser, é virtualmente não correlacionado com o índice do nível de recursos. Apesar de admitir que na eficiência de transformação existiam variações importantes, eles notam que elas em nada contribuíam para explicar a maior desigualdade no índice de bem-estar.

A estimação econométrica de funções de distância, para fins de análise de eficiência é bem recente, nada trivial e bem pouco explorada. A única pesquisa com uma função de distância de fronteira estocástica encontrada na literatura, é devida a Grosskopf & Hayes (1993). Elas estimam uma função de distância-insumo, através de um sistema de fronteira com equações de preços-sombra da forma

$$\begin{aligned}
D &= \sum_i \sum_j a_{ij} X_i^{1/2} X_j^{1/2} + \sum_j \sum_k b_{jk} X_j \ln Y_k \\
w_i / C &= a_{ii} + \sum_j a_{ij} X_i^{1/2} X_j^{1/2} + \sum_k b_{ik} \ln Y_k
\end{aligned} \tag{2.57}$$

onde: X_i é a quantidade do insumo i ;
 w_i / C é o preço do i -ésimo insumo deflacionado pelo custo e
 Y_k é o k -ésimo produto.

A estimação de (2.57) tem dois problemas econométricos. A primeira dificuldade advém da necessidade de colocar $D = 1$ no lado esquerdo da equação primária, pois D não tem valores observados. Ao fazer isto considera-se que os produtos sejam exógenos e os insumos, endógenos. Desta forma, os insumos e a estrutura de erros ficam correlacionados. No outro problema, a impossibilidade de estimar o intercepto causa viés nas estimativas por OLS, como decorrência da invariabilidade do lado esquerdo da equação principal. Para resolver estes problemas, são utilizadas variáveis instrumentais e as equações são estimadas pelo método das regressões aparentemente não relacionadas (*seemingly unrelated regressions* - SUR).

No tocante à decomposição da estrutura dos resíduos, Grosskopf e Hayes aplicam a metodologia do modelo ALS, apenas se diferenciando na distribuição adotada para o componente indicativo da eficiência técnica. Com base na média da distribuição condicional de u_i dado \mathbf{e}_i , eles calculam as estimativas individuais de u_i , seguindo Jondrow et al. (1982) - ver seção 3.1.3.5. Para estimar v empregam

$$\hat{v} = \mathbf{e} - E(u|\mathbf{e}) \tag{2.58}$$

que é utilizado no cálculo da medida de eficiência técnica para cada observação, da seguinte forma:

$$EFF = \frac{1}{\left(D(Y, X; \hat{\mathbf{b}}) + \hat{v} \right)} \tag{2.59}$$

Com a finalidade de dispor de informações sobre a precisão das estimativas obtidas, os pesquisadores usam a técnica de reamostragem conhecida como *bootstrap*. Através desta técnica produzem não só estimativas pontuais dos preços- sombra e da

eficiência técnica para cada UTD, como também constroem testes de hipótese sobre a significância de tais estimativas.

A conclusão principal do estudo é a constatação de que as UTD's pesquisadas não minimizam seus custos e que a metade delas super-utiliza o fator trabalho com relação ao fator capital. E que, de acordo com o nível médio de ineficiência técnica apurada, caberia uma redução de 10% no nível de utilização dos insumos.

2.6 FRONTEIRA DETERMINÍSTICA DE PRODUÇÃO MÚLTIPLA

O título “fronteira de produção múltipla” é utilizado aqui para designar a fronteira originária da estimação direta de uma função de produção uniequacional com mais de uma variável dependente. Não há registro na literatura sobre a aplicação desta abordagem na estimação de fronteiras estocásticas, com o objetivo de mensurar a eficiência técnica. Sob o enfoque determinístico, seu uso é bem recente, devendo-se a primeira experiência a Ruggiero (1996). No referido trabalho, a preocupação básica do autor é apresentar solução para o problema da persistência de folgas em insumos individuais, mesmo após a UTD ter alcançado a eficiência técnica, de acordo com o índice de Farrell.

Em um primeiro estágio, Ruggiero identifica os distritos Farrell-eficientes e os Koopmans-eficientes³, através de fronteiras determinísticas e não-paramétricas determinadas via DEA. Para isto, utiliza uma amostra de 3 produtos (serviços), 6 tipos de insumos e uma variável ambiental (indicador sócio-econômico) referentes a 636 escolas localizadas no estado de Nova York e ao ano escolar 1990-91.

De forma geral, ele modela implicitamente a produção como

$$f(Y) \equiv h(X|Z) \quad (2.60)$$

onde: $Y \equiv (y_1, \dots, y_s)$ é o vetor de serviços;

$X \equiv (x_1, \dots, x_m)$ é o vetor de insumos e

³ Segundo Ruggiero, uma UTD é Koopmans-eficiente quando é impossível aumentar qualquer produto (e/ou reduzir qualquer insumo), sem simultaneamente reduzir outro produto (e/ou aumentar algum outro insumo).

$Z \equiv (z_1, \dots, z_r)$ é o vetor de fatores ambientais.

Admitindo que a produção é tecnicamente eficiente, isto é, que a equação (2.60) define a produção, o pesquisador especifica como modelo de análise a seguinte função Cobb-Douglas estendida:

$$\sum_{j=1}^s a_j \ln y_j = \sum_{k=1}^m b_k \ln x_k + \sum_{l=1}^r g_l \ln z_l + e \quad (2.61)$$

onde: y_j é o j -ésimo produto;
 x_k é o k -ésimo insumo;
 z_l é o l -ésimo fator ambiental;
 e é um termo residual e
 a_j , b_k e g_l são os parâmetros a estimar.

Ruggiero estima (2.61) por CCA, para o que cria duas variáveis U e V , dadas por

$$U = a_1 \ln y_1 + \dots + a_s \ln y_s \quad (2.62)$$

$$e \quad V = b_1 \ln x_1 + \dots + b_m \ln x_m + c_1 \ln z_1 + \dots + c_r \ln z_r \quad (2.63)$$

Os pesos $A \equiv (a_1, \dots, a_s)$, $B \equiv (b_1, \dots, b_m)$, e $C \equiv (c_1, \dots, c_r)$ são escolhidos de forma a maximizar a correlação entre U e V . Assim,

$$\mathbf{r}^* = \max_{A,B,C} \text{Corr}(U, V) \quad (2.64)$$

leva às estimativas $A^* \equiv (a_1^*, \dots, a_s^*)$, $B^* \equiv (b_1^*, \dots, b_m^*)$, e $C^* \equiv (c_1^*, \dots, c_r^*)$. Como

$$U = \mathbf{r}^* V \quad (2.65)$$

uma estimativa de (2.61) é obtida pela substituição de (2.62) e (2.63) em (2.64), ou seja,

$$\sum_{j=1}^s a_j^* \ln y_j = \sum_{k=1}^m \mathbf{r}^* b_k^* \ln x_k + \sum_{l=1}^r \mathbf{r}^* c_l^* \ln z_l \quad (2.66)$$

Conforme salienta o pesquisador em tela, apoiado em Vinod (1968), a técnica de regressão canônica é de máxima verossimilhança (assumindo normalidade) e produz estimativas consistentes dos parâmetros dos regressores (mesmo sem assumir normalidade). As referidas estimativas são não-tendenciosas se e capta a ineficiência técnica e é não correlacionado com as variáveis independentes. Para evitar problema de viés pela violação da hipótese de eficiência, Ruggiero realiza a regressão canônica também

em uma sub-amostra de 208 escolas, identificadas como Koopmans-eficientes na análise DEA. Os resultados são similares aos obtidos para a regressão com a amostra completa, o que não surpreende o autor pelo fato de as escolas da amostra restrita terem sido utilizadas como referência na determinação das escolas ineficientes.

A equação estimada (2.66) propicia os pesos necessários para ajustar a medida de Farrell e acomodar as folgas existentes nos insumos. Para isto, Ruggiero constrói um índice derivado da razão entre o produto agregado observado (medido em termos de quantidade de insumo Koopmans-eficiente) e o produto agregado Koopmans-eficiente previsto (medido em termos da quantidade observada de insumo). O referido índice, similar em espírito à medida de Farrell (obtida via DEA), é dado por

$$y_i = \prod_{k=1}^m \left(\frac{x_{ik}^{\bullet}}{x_{ik}} \right)^{r^* b_k^*} \quad (2.67)$$

Com este procedimento, a medida DEA de ineficiência de Farrell é convertida em uma medida mais consistente com a definição de Koopmans. Ao invés de focar a redução equiproporcional, a medida pondera as razões de insumos pela contribuição de cada um na obtenção do produto agregado.

Este índice satisfaz as propriedades desejadas para as medidas de eficiência. Uma UTD é Koopmans-eficiente, se e somente se $y_i = 1$, pois $x_{ik}^{\bullet} = x_{ik}$ somente para os distritos Koopmans-eficientes. Para os estabelecimentos Koopmans-ineficientes, $x_{ik}^{\bullet} < x_{ik}$, o que implica em $y_i < 1$. Além disso, Ruggiero mostra que seu novo índice atende, também, as propriedades de homogeneidade e de monotonicidade.

A partir dos resultados da estimação, o autor calcula e interpreta as elasticidades marginais de cada produto com relação a cada insumo e a cada variável ambiental considerada, bem como as taxas marginais de transformação entre os produtos.

A conclusão principal da análise é a confirmação da hipótese de que a ineficiência das escolas seria a causa do persistente entendimento de que não havia muita diferença entre elas.

A partir do estudo de Ruggiero vem aumentando o interesse dos pesquisadores pela utilização da regressão canônica na mensuração da eficiência. Esta nova linha de pesquisa, à exceção do trabalho de Arnold et al. (1996) mostrado na seção 2.2, tem se concentrado no uso conjunto das técnicas CCA/DEA. Assim procedem Friedman & Stern (1997) que chamam de um “novo método” a utilização de CCA para estabelecer um *ranking* dos escores de eficiência para todas as UTD's, ao invés de uma classificação em unidades eficientes e ineficientes como é feito por DEA. Em uma aplicação, Tofallis (1996) usa o método híbrido para modelar a fronteira tecnológica de produção e, em outra, utiliza a mesma amostra e os mesmos procedimentos para sugerir uma nova medida de eficiência da produção (Tofallis, 1998).

Em fins desse mesmo ano, Ruggiero (1998) volta à cena com um novo enfoque, em que especifica a função de produção Cobb-Douglas, na forma estendida,

$$y_1^{a_1} \dots y_s^{a_s} \equiv g x_1^{b_1} \dots x_m^{b_m} \quad (2.68)$$

onde: $y_1 \dots y_s$ são os s produtos;

$x_1 \dots x_m$ são os m insumos;

$a_1 \dots a_s$ e $b_1 \dots b_m$ são os parâmetros a estimar e

g é um índice de eficiência ($0 < g \leq 1$).

Ruggiero estima a forma logarítmica de (2.68) por CCA, produzindo

$$\sum_{j=1}^s a_j^* \ln y_j = \sum_{k=1}^m r^* b_k^* \ln x_k + \ln g \quad (2.69)$$

onde a_j^* , r^* e b_k^* têm os mesmos significados já definidos para (2.66).

Com base em (2.69), o pesquisador constrói um modo alternativo para estimar a eficiência técnica de UTD's com múltiplos produtos:

$$\ln g = \sum_{j=1}^s a_j^* \ln y_j - \sum_{k=1}^m r^* b_k^* \ln x_k \quad (2.70)$$

de onde uma estimativa da eficiência de Farrell é derivada como

$$\hat{g} = \exp \left[\sum_{j=1}^s a_j^* \ln y_j - \sum_{k=1}^m r^* b_k^* \ln x_k \right] \quad (2.71)$$

Para fins de avaliação do modelo proposto, o autor usa uma matriz de dados com 300 observações geradas aleatoriamente por intermédio do SAS. Na simulação, são geradas 89 observações (UTD's) como tecnicamente eficientes. São considerados três insumos (com distribuições uniformes) e dois produtos, relacionados de acordo com a seguinte função de produção:

$$y_1^{0,45} y_2^{0,55} = g x_1^{1/3} x_2^{1/3} x_3^{1/3} \quad (2.72)$$

Além destas 5 variáveis, são também geradas x_4 , x_5 e y_3 que não são relevantes para a produção. Este procedimento serve para avaliar o impacto decorrente da introdução de variáveis estranhas.

A eficiência técnica é então, mensurada usando DEA e CCA e os respectivos índices são comparados com os escores de eficiência verdadeiros (gerados pela simulação).

A análise comparativa é levada a efeito através dos coeficientes de correlação ordinal de Pearson, de Spearman e de Kendall e do teste de Wilcoxon de pares combinados em uma amostra. Neste teste é investigada a hipótese nula de que a diferença absoluta (Índice CCA - Índice verdadeiro) seja menor do que a diferença absoluta (Índice DEA - Índice verdadeiro).

Com os mesmos procedimentos empregados para os dados simulados, Ruggiero realiza a análise empírica da produção educacional usando dados sobre os distritos escolares do estado de Nova York no ano letivo 1990-91.

Em suas conclusões, Ruggiero ressalta que seu novo método é um forte concorrente para a análise DEA. Como possibilidade para pesquisas futuras, sugere testar a nova técnica quanto à especificação da forma funcional e tentar estender o enfoque CCA para abrigar a inclusão de um componente estocástico.

2.7 SÍNTESE E CONSIDERAÇÕES

O uso de modelos de fronteira vem sendo crescentemente difundido por uma variedade de razões. Dentre elas, ressalta o fato da noção de fronteira ser consistente com

a teoria econômica básica, quanto à otimização. Outra razão importante, é a interpretação natural dos desvios até a fronteira como uma medida de eficiência, segundo a qual as UTD's perseguem suas metas técnicas, dentre outras. Também, merece menção a sua larga utilidade e aplicação na formulação de políticas com base na estrutura da fronteira e na eficiência relativa das unidades econômicas.

O modelo de fronteira estocástica de produção foi criado de forma independente, mas simultânea, por Aigner et al. (1977), Meeusen & van den Broeck (1977) e Battese & Corra (1977). Desde então, a formulação mais referenciada é a do primeiro grupo, a qual é um modelo a uma equação com somente uma variável dependente representando a produção. Esta abordagem constitui-se, até fins da década de 80, na única alternativa de tratamento uniequacional de fronteiras com múltiplos produtos. Sob o enfoque multi-produto, na concepção de Lovell (1995), a agregação dos vários produtos deve ser feita por meio de um índice como o empregado por Gathon & Pestieau (1995). Mais tarde, Arnold et al. (1996) utilizam os coeficientes canônicos, oriundos da estimação por CCA, na tarefa de agregação. Para Lovell, a fronteira de produção não acomoda desagregadamente múltiplos produtos, como as fronteiras de custo, de receita e de lucro. Segundo este pesquisador, a abordagem desagregada em uma fronteira de produção só pode ser viabilizada indiretamente, através da estimação de uma fronteira de necessidades de insumo (ou de produção inversa) e pela estimação de uma fronteira de distância-produto. Deste modo, segundo a classificação de Lovell (1995) para a abordagem de múltiplos produtos, existem na análise econométrica de eficiência os seguintes modelos uniequacionais: fronteira de valor dual, fronteira de produção agregada, fronteira de produção inversa e fronteira de distância-produto.

Os aspectos principais referentes à especificação e aos resultados empíricos, levantados na literatura para cada um desses modelos, integram as seções anteriores do presente capítulo. Nesta seção, são identificadas e avaliadas as especificações mais consagradas que sejam de interesse para a presente pesquisa.

Existem dois paradigmas sobre como construir fronteiras. Um deles, a análise DEA, usa técnicas de programação matemática e, o outro, emprega técnicas

econométricas. Enquanto o enfoque DEA é não-paramétrico, não-estocástico e manuseia facilmente com tecnologias de produção multi-produto, a análise econométrica é paramétrica, estocástica e enfrenta dificuldades para tratar fronteiras com múltiplos produtos. Por conseguinte, o uso conjunto DEA/econometria visando a comparação dos resultados depende do equacionamento destas diferenças, primeiramente no âmbito próprio de cada método. Para isto, muita pesquisa vem sendo realizada, com destaque para o desenvolvimento do modelo DEA/estocástico, na análise DEA, e para o emprego de sistemas de equações de demanda, na análise econométrica. Mesmo assim, a abordagem da presente pesquisa, apesar de não tão avançada quanto o tratamento multiequacional, não é menos importante no que diz respeito ao encaminhamento da problemática da produção múltipla em econometria.

Um assunto recorrente na estimação de fronteiras tem sido o conflito entre estrutura e flexibilidade. Idealizadamente, este conflito é evitado quando a estrutura correta é conhecida e imposta *a priori*, ou quando é estimado um modelo suficientemente flexível, tal que as restrições possíveis possam ser testadas. Por isso, a estrutura apropriada deve ser determinada mediante cuidadosas considerações a respeito dos dados e das características das UTD's sob estudo. Todavia, nem sempre existem testes estatísticos para orientar o caminho e o pesquisador tem deixado de justificar as suas escolhas dentre as especificações possíveis. Mesmo em situações em que existem critérios bem definidos para subsidiar a escolha, não se tem notado muita preocupação com a sua justificativa.

Na especificação da forma funcional e do número de equações do modelo de fronteira, é reconhecido que as formas funcionais flexíveis abrangem tecnologias mais sofisticadas e que a estimação de sistemas de equações leva a estimativas assintoticamente mais eficientes. No entanto, a opção tem recaído, preferencialmente, em funções a uma equação do tipo Cobb-Douglas. Ir além de uma função de fronteira Cobb-Douglas uniequacional, significa muito mais dificuldades e trabalho na tarefa de estimação. No caso do emprego de um sistema de equações, o analista se depara com um difícil problema ainda não completamente resolvido chamado "*Greene problem*" (ver detalhes no final da seção 2.3)

Usualmente têm sido empregadas fronteiras estocásticas porque elas consideram o ruído estatístico resultante de eventos que estão fora do controle da firma, além do efeito de erros de especificação e de omissão de variáveis não correlacionadas com os regressores incluídos. Por conseguinte, obter estimativas de eficiência relativa para cada UTD, tem muito mais a ver com uma fronteira estocástica do que determinística, pois, nesta última, os índices são produzidos diretamente dos resíduos da estimação.

Em modelos de fronteira estocástica, a ineficiência é modelada por um termo residual integrado pelo ruído estatístico e por uma perturbação unicaudal indicativa do nível de ineficiência. A menos que seja usado um painel de dados, é preciso definir a distribuição do componente unilateral. Para fazer isto os pesquisadores têm sido guiados pela conveniência e pela intuição, devido à falta de um maior direcionamento da teoria econômica. A maioria dos analistas tem optado pela distribuição seminormal, porém são raros os que fazem testes estatísticos de aderência para conferir a sua decisão, ou para tomá-la.

Os dois métodos comumente utilizados na estimação de funções de fronteira estocástica são MLE e COLS, com nítida preferência para o primeiro. A exígua presença dos estimadores COLS na literatura aplicada é atribuída, em grande parte, à falta de acesso aos seus erros-padrão (Coelli, 1995). A estimação pelo método CCA é encontrada apenas no caso de fronteiras determinísticas, em que é empregado em conjunto com o método DEA. Não há nenhum registro sobre o uso do método CCA para estimar a fronteira estocástica.

No caso de fronteiras duais, Lovell (1993 e 1995) menciona apenas uma aplicação da fronteira de custo (Hughes, 1988), a qual trata da mensuração da eficiência produtiva. A decomposição desta medida, nos componentes referentes à eficiência técnica e à eficiência alocativa, enfrenta enormes problemas conforme se pode constatar no final da seção 2.3. Por isso, naquela seção são incluídas duas outras aplicações em que as fronteiras de custo uniequacionais satisfazem as condições necessárias à estimação somente da eficiência técnica.

A obtenção direta unicamente dos escores de eficiência técnica, por meio da estimação de fronteiras de lucro e de receita a uma equação, ainda carece de desenvolvimento teórico. Kumbhakar (1996) trabalha com uma fronteira de lucro, sob a hipótese de maximização de lucros, porém não oferece a sua forma estimável. Por isso, a presente pesquisa não inclui as fronteiras de lucro e de receita. Esta última é contemplada indiretamente com a inclusão de sua fronteira primal, que é a fronteira de distância-produto.

METODOLOGIA

No capítulo 2 objetiva-se levantar e conhecer o estado-da-arte dos principais aspectos metodológicos e resultados referentes à especificação e à estimação de modelos econométricos de fronteira a uma equação.

No presente capítulo, os referidos aspectos são detalhados e especificados de forma a constituir o corpo metodológico próprio da pesquisa.

O alicerce da construção metodológica da pesquisa é apoiado nos fundamentos teóricos sobre a representação da tecnologia de produção e a mensuração do grau de eficiência técnica com que a tecnologia é operada. Esta formalização passa pelas funções de fronteira e pelos métodos disponíveis para estimá-las, dentre os quais se insere a proposição original da pesquisa, desembocando nos procedimentos de estimação da eficiência técnica para cada UTD.

Com base nessa fundamentação teórica e nos aspectos metodológicos levantados na revisão da literatura, são definidos os modelos de análise da pesquisa. Primeiramente são especificados todos os modelos vigentes e, por último, a proposição do novo modelo designado “fronteira de produção múltipla”. A tarefa de especificação dos modelos de análise leva em conta que o objetivo máximo da pesquisa é desenvolver e apresentar um novo modelo, sob condições bem definidas. Sob esta ótica, são especificados, preferencialmente, os requisitos destinados a garantir condições satisfatórias de comparabilidade dos resultados.

Para atender os referidos requisitos, as especificações, em algumas situações, envolvem procedimentos não-triviais. É o caso da estimativa dos resíduos indicativos da ineficiência técnica, do uso da técnica *jackknife* para estimar os erros-padrão dos coeficientes canônicos e da verificação das condições de primeira e de segunda ordens referentes à função de fronteira de produção múltipla.

Após a especificação dos modelos de análise, são definidos os critérios para a construção dos dados amostrais e para a avaliação comparativa dos resultados. A primeira tarefa inclui, também, conduta não-trivial, uma vez que a geração das amostras aleatórias é seguida pela estimação e verificação de cada modelo de análise. Já a avaliação dos resultados, não precisa ir além do cotejamento dos índices de eficiência, em termos de correlação e de variabilidade.

A última seção trata a respeito da elaboração dos programas computacionais utilizados no processamento dos dados da pesquisa.

3.1 FUNDAMENTOS TEÓRICOS

3.1.1 Tecnologia de produção e eficiência técnica

O produtor, ou a unidade produtora, ou a firma, ou a indústria, seja de caráter privado ou público, nesta pesquisa chamada de UTD, é um agente econômico que utiliza um conjunto de insumos (fatores de produção) e os transforma (ou gerencia) produzindo um conjunto de produtos (bens e/ou serviços). Em outras palavras, o produtor utiliza uma

certa combinação de insumos, principalmente mão-de-obra e capital, para produzir certas combinações de produtos. Então, cada nível de produto está associado a diferentes combinações de insumos. Um conjunto de insumos, que pode ser usado para produzir um dado nível de produtos, é dito ser viável (ou possível) para aquele nível. Reciprocamente, cada combinação de insumos está associada a diferentes níveis de produto. Esta complexa relação entre insumos e produtos é denominada “tecnologia de produção”. Neste sentido, a tecnologia é interpretada como uma coleção de normas técnicas, métodos e fórmulas que permitem aos produtores determinar que combinação de insumos pode ser usada para obter um dado nível de produto. Por isso, a tecnologia é a referência mais natural ou padrão utilizada para definir a eficiência de um modo formal.

A produção, por sua vez, é definida como um processo de transformação e/ou gerenciamento de um conjunto de insumos $X = (x_1, \dots, x_m) \in \mathfrak{R}_+^m$ para produzir um conjunto de produtos $Y = (y_1, \dots, y_p) \in \mathfrak{R}_+^p$. Este processo de transformação e/ou gerenciamento ocorre de acordo com uma dada tecnologia de produção, que é representada formalmente pelo conjunto das necessidades de insumo

$$L(Y) = \{X: (Y, X) \text{ é viável produzir} \} \quad (3.1)$$

Esta relação indica que, para o vetor de produtos Y , qualquer integrante do conjunto das necessidades de insumo é suficiente para produzir o referido vetor. Neste conjunto, ainda não se faz nenhuma menção à questão da eficiência, nem se define a função de produção. Isto é feito, indiretamente, na medida em que se especifica o conjunto de insumos que é insuficiente para produzir Y (o complemento de $L(Y)$ em \mathfrak{R}_+^m) e, portanto, demarca-se os limites da capacidade de produção. Assim, a função de produção é estabelecida pela isoquanta

$$\text{Isoq } L(Y) = \{X: X \in L(Y), \quad \lambda X \notin L(Y), \quad \lambda \in (0,1)\} \quad (3.2)$$

que determina a fronteira do conjunto das necessidades de insumo.

Uma outra representação da tecnologia de produção, introduzida por Shephard *apud* LOVELL (1993), é a função de distância-insumo

$$D_{\text{inp}}(Y, X) = \max \left\{ \tilde{\lambda}: \left[\frac{1}{\tilde{\lambda}} \right] X \in L(Y) \right\} \quad (3.3)$$

É claro que

$$D_{\text{inp}}(Y, X) \geq 1 \quad (3.4)$$

e que Isoq $L(Y)$ é o conjunto dos m insumos para o qual

$$D_{\text{inp}}(Y, X) = 1 \quad (3.5)$$

Uma importante propriedade desta função é a sua homogeneidade de grau 1 nos insumos, ou seja,

$$D_{\text{inp}}(Y, \tilde{\lambda} X) = \tilde{\lambda} D_{\text{inp}}(Y, X) \quad \forall (X, Y) \in \mathfrak{R}_+^{m+p}, \quad \tilde{\lambda} > 0 \quad (3.6)$$

Esta propriedade, associada às exigências de convexidade de $L(Y)$ e de descartabilidade forte dos insumos, é suficiente para garantir a dualidade entre a função de distância-insumo e a função de custo.

O ponto natural de partida para a análise de eficiência é a medida insumo-orientada de eficiência técnica, devida a Debreu (1951) e a Farrell (1957),

$$DF_{\text{inp}}(Y, X) = \min \{ \tilde{\lambda}: \tilde{\lambda} X \in L(Y) \} \quad (3.7)$$

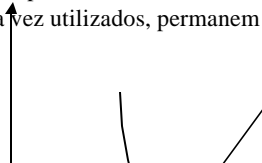
em que, por definição,

$$DF_{\text{inp}}(Y, X) \leq 1$$

$$\text{e} \quad DF_{\text{inp}}(Y, X) = \frac{1}{D_{\text{inp}}(Y, X)} \quad (3.8)$$

A medida de Debreu-Farrell é estritamente expressada em termos de produção e é, conforme considerada aqui, uma medida de eficiência técnica ligada à contração ou expansão radial do vetor de insumos. A figura 3.1 mostra o gráfico convencional de uma isoquanta/isocusto para um único produto, que é produzido com o uso de dois tipos de insumo e cuja razão de preços é determinada pela inclinação da linha de isocusto $(c_0 = w_1 x_1 + w_2 x_2)$.

⁴ Daqui em diante são utilizadas letras maiúsculas para representar as matrizes/vetores de variáveis e letras minúsculas para as variáveis individuais (escalares). Cada letra e símbolo, incluindo os indexadores, uma vez utilizados, permanem com o mesmo significado até o final.



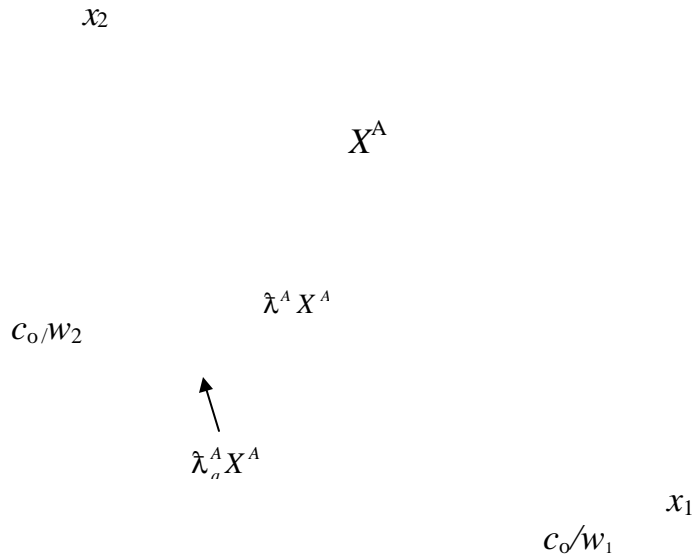


Figura 3.1 - Gráfico convencional de uma isoquanta/isocusto.

A medida de Debreu-Farrell de eficiência técnica é dada por $\hat{\lambda}$ que, em termos econômicos, não enfatiza o grau de eficiência alocativa. Reduzindo ambos os insumos na mesma proporção $\hat{\lambda}$, o produtor A alcançaria a isoquanta de produção e, assim, passaria a ser tecnicamente eficiente. Porém, poderia estar sendo alocativamente ineficiente, pois poderia estar produzindo a um custo mais baixo.

Até aqui, fez-se uso da representação da tecnologia no conjunto das necessidades de insumo $L(Y)$. Quando a tecnologia de produção é representada por um conjunto de produção, tem-se

$$P(X) = \{Y: (X, Y) \text{ é possível} \} \quad (3.9)$$

o qual, para cada $X \in \mathfrak{R}_+^m$, é um conjunto limitado, fechado e convexo, que tem isoquanta

$$\text{Isoq } P(X) = \{Y: Y \in P(X), \mathbf{d}Y \notin P(X), \mathbf{d} \in (1, +\infty)\} \quad (3.10)$$

e subconjunto eficiente

$$\text{Eff } P(X) = \{Y: Y \in P(X), Y' \notin P(X), Y' \geq Y\} \quad (3.11)$$

com a propriedade

$$\text{Eff } P(X) \subseteq \text{Isoq } P(X) \quad (3.12)$$

A representação paramétrica da tecnologia de produção, através de uma função matemática, impõe igualdade entre as isoquantas e os sub-conjuntos eficientes, suprimindo a presença de folgas. Por isso, a diferença entre o índice de Debreu-Farrell e o de Koopmans (1951), que considera tecnicamente eficientes apenas os produtores sediados sobre o sub-conjunto eficiente, não apresenta problemas na análise econométrica de eficiência.

Uma outra representação para $P(X)$, idealizada simultaneamente com a função de distância insumo-orientada (3.3), pelo mesmo pesquisador, é a função de distância-produto.

$$D_{\text{out}}(X, Y) = \min \{ \mathbf{d} : (Y / \mathbf{d}) \in P(X) \} \quad (3.13)$$

Ambas as representações $D_{\text{inp}}(X, Y)$ e $D_{\text{out}}(X, Y)$, oferecem tratamentos funcionais alternativos para tecnologias com múltiplos produtos.

Se um dado vetor X determina o conjunto de produção $P(X)$, e um vetor de produtos Y^A é arbitrariamente escolhido, o valor de $D_{\text{out}}(X, Y^A)$ coloca $\frac{Y^A}{D_{\text{out}}(X, Y^A)}$ na fronteira de $P(X)$ e sobre o raio através de Y^A . Nestas condições, desenhadas na figura 3.2 adiante, Y^A está no interior de $P(X)$ e, por isso,

$$D_{\text{out}}(X, Y^A) \leq 1 \quad (3.14)$$

$$\text{ou, por (3.10),} \quad \text{Isoq } P(X) = \{ Y : D_{\text{out}}(X, Y) = 1 \} \quad (3.15)$$

De acordo com Färe & Primont (1995), quando $D_{\text{out}}(X, Y)$ é bem definida, ou seja, $P(X) \subseteq P_{D_{\text{out}}}(X)$, ela é semi-limitada e semi-contínua, inferiormente, em \Re_+^p , e sempre homogênea de grau 1 nos produtos, isto é, respectivamente,

$$\{ Y : D_{\text{out}}(X, Y) \leq \mathbf{j} \} \text{ é limitado } \forall \mathbf{j} > 0 \text{ e } X \in \Re_+^m \quad (3.16)$$

$$\{ Y : D_{\text{out}}(X, Y) \leq \mathbf{j} \} \text{ é fechado } \forall \mathbf{j} > 0 \text{ e } X \in \Re_+^m \quad (3.17)$$

$$\text{e} \quad D_{\text{out}}(X, \mathbf{v} Y) = \mathbf{v} D_{\text{out}}(X, Y), \forall \mathbf{v} > 0 \text{ e } \forall (X, Y) \in \Re_+^{m+p} \quad (3.18)$$

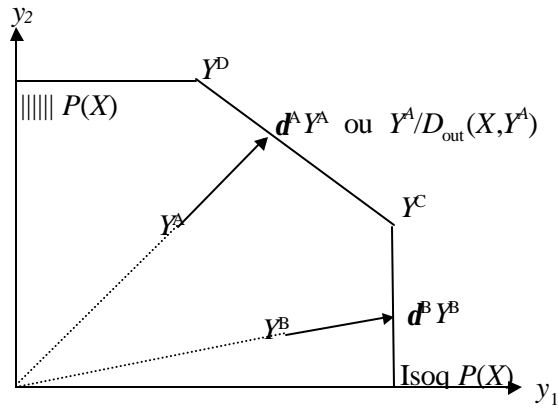


Figura 3.2 - Medida de eficiência técnica de Debreu-Farrell orientada para o produto.

A medida de eficiência técnica produto-orientada de Debreu-Farrell, representada na figura 3.2, é definida por

$$DF_{out}(X, Y) = \max \{d: dY \in P(X)\} \quad (3.19)$$

$DF_{out}(X, Y) \geq 1$, sendo que, por (3.13), tem-se

$$DF_{out}(X, Y) = \frac{1}{D_{out}(X, Y)} \quad (3.20)$$

$$\text{e, conseqüentemente, } Isoq P(X) = \{Y: DF_{out}(X, Y) = 1\} \quad (3.21)$$

Lovell (1993) sintetiza as principais propriedades das medidas de eficiência técnica de Debreu-Farrell, através dos três tópicos seguintes:

- 1) $DF_{inp}(X, Y)$ é homogênea de grau 1 nos insumos e $DF_{out}(X, Y)$ é homogênea de grau 1 nos produtos;
- 2) $DF_{inp}(X, Y)$ é monotônica e levemente decrescente nos insumos e $DF_{out}(X, Y)$ é monotônica e levemente decrescente nos produtos e
- 3) $DF_{inp}(X, Y)$ e $DF_{out}(X, Y)$ não variam com as mudanças nas unidades de medida das variáveis consideradas.

3.1.2 Funções de produção de fronteira determinística e eficiência técnica

O enfoque neoclássico da teoria da produção também descreve a tecnologia através de um determinado conjunto $L(Y)$, definido em um espaço real n -dimensional, associado a cada nível de produto Y e chamado conjunto-insumo. Sob algumas suposições

impostas axiomáticamente, esta descrição pode ser sintetizada com a ajuda de uma função $f(X)$ denominada “função de produção neoclássica”. As suposições impostas sobre os conjuntos-insumo garantem a existência desta função, com algumas propriedades e hipóteses que estão presentes na análise econômica desde antes do neoclassicismo. Assim, a questão do conjunto-insumo $L(Y)$ pode ser caracterizada pela desigualdade $y \leq f(X)$.

Esta inequação, por sua vez, conduz naturalmente à denominação “fronteira de produção” da função de produção neoclássica. Deve-se perceber que, alternativamente, uma função f mapeando o espaço dos insumos sobre o espaço dos produtos, pode ser definida com propriedades axiomáticamente estabelecidas. Esta função implica na existência de uma família de conjuntos-insumo $L(Y)$.

As idéias que estão por trás do conceito de fronteira de produção são bastante antigas e aparecem bem cedo na história da economia. Entretanto, a representação rigorosa e formal do processo de produção (tecnologia) é mais recente. Ela introduz a função de produção como uma fronteira limitante de um conjunto de possibilidades de produção. Este aspecto deve ficar bem realçado. O desenvolvimento teórico de funções de produção trabalha com a idéia de uma fronteira de produção que indica o máximo de produto a ser obtido com um dado montante de insumos. Quando se compara o resultado real (observado) do processo produtivo com o resultado padrão dado pela fronteira, chega-se naturalmente ao problema de mensuração da eficiência produtiva, que engloba a eficiência técnica e a eficiência alocativa.

A mensuração empírica da eficiência técnica requer a definição de uma função de transformação, ou seja, de uma “função de produção de fronteira”. Sendo

$$y = fpf(X) \quad (3.22)$$

esta função para produzir um único produto y , usando um vetor de insumos X , a medida de eficiência técnica de Debreu-Farrell orientada para o produto $DF_{out}(X,Y)$, doravante designada por $ET(X,Y)$, é dada por

$$ET(Y, X) = \frac{y}{fpf(X)} \quad (3.23)$$

que é a medida convencional da produtividade total dos fatores.

Esta medida é incorporada pelo método econométrico bem de acordo com a definição dos livros-texto de teoria da produção, sendo modelada por

$$y_i = f_{pf}(X_i, \phi) ET_i \quad (3.24)$$

$$\text{onde: } 0 < ET(X_i, y_i) \leq 1; \quad (3.25)$$

ϕ é o vetor de parâmetros a ser estimado e

i indexa as n UTD's da amostra.

Quando o modelo de produção é linear nos logaritmos das variáveis, a equação (3.24) fica sendo da forma

$$\begin{aligned} \ln y_i &= \ln f_{pf}(X_i, \phi) + \ln ET_i \\ &= \ln f_{pf}(X_i, \phi) - u_i \end{aligned} \quad (3.26)$$

onde $u_i > 0$ é a medida de eficiência técnica, uma vez que

$$u_i = -\ln ET_i \approx 1 - ET_i \quad (3.27)$$

$$\text{e} \quad ET_i = e^{-u_i} \quad (3.28)$$

Assim sendo, depreende-se que a mensuração econométrica da medida de eficiência técnica de Debreu-Farrell deve ser precedida da estimação de uma função de produção de fronteira.

As primeiras estimações de fronteiras de produção são de funções de produção “médias”, com a finalidade de subsidiar a pesquisa metodológica de novas formas funcionais. Essas funções pressupõem que as firmas produzem no limite de seus conjuntos de possibilidades tecnológicas, ou perto dele, ficando as diferenças simétrica e estocasticamente distribuídas em ambos os lados da fronteira. Este tipo de estimação não possibilita a mensuração da eficiência técnica como preconizada por Farrell.

As primeiras estimações de fronteiras de produção, materializando a noção de uma fronteira de produção máxima, devem-se a Aigner & Chu (1968). Eles estimam funções de produção tipo Cobb-Douglas

$$y_i = \mathbf{b} \cdot \prod_{k=1}^m x_{ki}^{b_k} u_i \quad (3.29)$$

onde: \mathbf{b} e b_k são os parâmetros a estimar e

u_i (que corresponde a ET_i) é uma perturbação aleatória contida entre 0

e 1.

Tomando os logaritmos das variáveis, tem-se

$$\begin{aligned}\ln y_i &= \mathbf{b}_0 + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \mathbf{e}_i \\ &= \mathbf{b}_0 + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} - u_i\end{aligned}\tag{3.30}$$

onde $\mathbf{b}_0 = \ln \mathbf{b}_\bullet$ e $\mathbf{e}_i = \ln u_i$.

Neste enfoque, a parte não-estocástica, que fica do lado direito da equação, é interpretada como sendo a fronteira. Fronteiras deste tipo são chamadas determinísticas, porque o componente estocástico está inteiramente inserido no termo $-u_i$, ou seja, os desvios de cada observação de produção até o seu máximo teórico na fronteira (ideal, potencial) são atribuídos unicamente à ineficiência da UTD sob análise.

A estimação de funções de fronteira está assentada nas suposições de que \mathbf{e}_i é aleatoriamente distribuído entre as n UTD's envolvidas, tem distribuição independente de todas as variáveis do modelo, média diferente de zero, variância constante e diferentes resíduos de regressão ordinária. Estas suposições pressupõem $E[\mathbf{e}_i] < 0$ o que, na ausência de quaisquer outras considerações especiais, mantém (3.30) como um modelo clássico de regressão linear, que também pode ser escrito como

$$\ln y_i = (\mathbf{b}_0 + E[\mathbf{e}_i]) + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + (\mathbf{e}_i - E[\mathbf{e}_i])\tag{3.31}$$

Esta especificação continua sendo a de um modelo clássico de regressão linear. A normalidade da perturbação está obstada, uma vez que $(\mathbf{e}_i - E[\mathbf{e}_i])$ é a diferença entre uma variável aleatória, que é sempre negativa, e a sua média. Apesar disso, os parâmetros do modelo podem ser estimados consistentemente por OLS, já que estes estimadores são robustos para não-normalidade. Deste modo, os parâmetros técnicos de produção, com exceção do coeficiente de interseção, podem ser estimados consistentemente, embora não eficientemente, por OLS⁵. Portanto, neste contexto, o coeficiente de interseção é o principal parâmetro de interesse (Greene, 1993a).

⁵ O termo residual é distribuído não-normalmente por construção, de modo que o estimador de máxima verossimilhança é mais eficiente.

Assim, nos modelos de fronteira determinística, os parâmetros dos regressores podem ser estimados consistentemente por OLS. O coeficiente de interseção, por sua vez, pode ser estimados consistentemente através do deslocamento da linha ou do plano de regressão suficientemente para cima, de tal modo que o maior resíduo seja igual a zero. Em outras palavras: até o ponto em que todos os resíduos, exceto um, no qual se pendura a função, sejam negativos. Este método é conhecido como método COLS e implica na medida de eficiência

$$-u_i = e_{i\text{cols}} - \max_i e_{i\text{cols}} \quad (3.32)$$

onde $e_{i\text{cols}}$ são os resíduos de mínimos quadrados corrigido, ou seja, calculados após a correção do coeficiente de interseção obtido por OLS.

Um enfoque alternativo, é o do método MOLS que requer um modelo paramétrico da distribuição de u_i , muitas vezes suposta ser exponencial ou gama. Este método é um pouco menos ortodoxo do que o método COLS, uma vez que permite que alguns resíduos sejam positivos.

Nos modelos de fronteira determinística, qualquer erro de medida e qualquer outra fonte de variação estocástica da variável dependente está embutida na distribuição unilateral. Como tal, em qualquer amostra, uma única observação exageradamente distorcida pode ter profundos efeitos sobre as estimativas. E isto é agravado pelo fato da ocorrência de *outliers* não recrudescer, quando se recorre a amostras maiores. Portanto, sob a interpretação de uma fronteira determinística, por exemplo, um não-habitual e elevado número de registros de falhas nos equipamentos, ou até mesmo o mau tempo, podem ser computados pelo analista como ineficiência da UTD.

Principalmente por estes motivos, a partir deste ponto deixa-se de lado as fronteiras determinísticas e passa-se a tratar, exclusivamente, dos aspectos teóricos relativos à especificação e à estimação de fronteiras não-determinísticas.

3.1.3 Funções de produção de fronteira estocástica e eficiência técnica

3.1.3.1 Introdução

O modelo econométrico de fronteira estocástica de produção, proposto por Aigner et al⁶ (1977), é inspirado na idéia de que os desvios da produção observada (efetivamente produzida) até a produção de fronteira (ideal), podem não estar inteiramente sob o controle da UTD.

No modelo de fronteira não-estocástica (ver final da seção 3.1.2), um não-habitual e alto número de equipamentos falhando aleatoriamente, ou mesmo a ocorrência de mau tempo, podem ser interpretadas pelo analista como ineficiência técnica. Pior ainda, qualquer erro e inadequação na especificação do modelo e na medição das variáveis envolvidas pode, também, resultar em uma medida de ineficiência distorcida. Estas razões desestimulam em muito o uso da fronteira determinística.

Uma interpretação mais razoável é a de que cada produtor faça frente à sua própria fronteira de produção, a qual é determinada aleatoriamente por toda uma coleção de elementos estocásticos, que podem entrar no modelo fora do controle do produtor. Assim, uma apropriada formulação é

$$y_i = fpf(X_i, \theta) e^{V_i + U_i} \quad (3.33)$$

cujas correspondente forma logaritmada é

$$\ln y_i = \ln fpf(X_i, \theta) + V_i + U_i \quad (3.34)$$

onde: V_i é o componente residual aleatório que captura os efeitos do ruído estatístico. É suposto ser independente e identicamente distribuído (iid) como uma $N(0, \sigma_V^2)$;

U_i é o componente residual que representa a ineficiência técnica. É suposto ser distribuído independentemente de V_i e satisfazer a condição $U_i \leq 0$;

$$\ln fpf(X_i, \theta) \quad (3.35)$$

é a fronteira determinística de produção e

$$\ln fpf(X_i, \theta) + V_i \quad (3.36)$$

é a fronteira estocástica de produção.

⁶ Este modelo está apresentado na seção 2.1 do capítulo 2, onde é abreviadamente denominado ALS.

De acordo com esta modelagem, Lovell (1993) considera que o grau de eficiência técnica de um produtor é dado pela razão entre o nível de produto observado e o seu correspondente nível na fronteira estocástica de produção, a qual é determinada pela fronteira determinística de produção e por eventos externos fora do controle do produtor. Para Lovell,

$$\ln ET_i = \ln y_i - [\ln fpf(X_i, \phi) + V_i] = U_i \quad (3.37)$$

é a versão econométrica da medida de eficiência técnica de Debreu-Farrell com base no produto, que é a recíproca de $DF_{out}(X, Y)$.

Greene (1993a) mostra que pelo fato de V_i e U_i serem geralmente admitidos como iid, a ineficiência média presente na distribuição fica refletida na sua assimetria. Ela é facilmente estimável (até) por OLS através do terceiro momento amostral dos resíduos, o qual é dado por

$$m_3 = \frac{1}{n} \sum_{i=1}^n (\hat{e}_i - \hat{E}[\mathbf{e}_i])^3 \quad (3.38)$$

Através da versão de (3.38) na população

$$\mathbf{m} = E[V_i - (U_i - E[U_i])]^3 \quad (3.39)$$

ele conclui que, de fato, a inclinação da distribuição de \mathbf{e}_i é simplesmente o negativo do seu componente de ineficiência latente, U_i . Por isso, mesmo sem levar em conta a distribuição básica definida, o negativo do terceiro momento dos resíduos de OLS oferece um estimador consistente da inclinação da distribuição de U_i .

A perturbação composta do modelo de fronteira estocástica, enquanto assimetricamente distribuída, é, para a grande maioria das distribuições de resíduos escolhidas, quanto ao mais, bem comportada. A distribuição da variável aleatória composta $\mathbf{e}_i = (V_i + U_i)$ está exaustivamente derivada em Weinstein (1964) e é amplamente discutida em Aigner et al. (1977).

Desde a proposição do modelo ALS, em 1977, tem surgido um considerável elenco de distribuições alternativas para U_i : a truncada normal, a seminormal, a exponencial, a gama/normal, a gama/exponencial e, mais recentemente, o enfoque Bayesiano e a forma Erlang do modelo normal-gama de Greene. Não obstante, a

especificação seminormal tem prevalecido na grande maioria dos estudos empíricos realizados. A distribuição seminormal pressupõe que muitas firmas estejam agrupadas perto da região de ineficiência zero e que, gradativamente, bem poucas firmas apresentem-se cada vez mais e mais ineficientes.

3.1.3.2 Estimação pelo método da máxima verossimilhança

A estimação de funções de fronteira estocástica por MLE, sob qualquer hipótese distribucional para U_i , é direta. Ela se concretiza quando é atingido o valor máximo da correspondente função logarítmica de máxima verossimilhança ($LogL$). Todas as estimativas produzidas por MLE são consistentes e assintoticamente eficientes (Aigner et al. 1977).

Os valores iniciais das iterações são tipicamente obtidos da manipulação dos resultados de OLS para chegar aos estimadores de momentos dos parâmetros da distribuição básica. Quando os U_i 's são seminormais, os valores iniciais das iterações são as estimativas deste modelo com a média dos U_i 's (m) igual a zero. Se este nível adicional de generalidade é justificado, ainda resta ser explorado. O benefício disto é, obviamente, o relaxamento de uma restrição possivelmente errada. E o custo inerente, advém do fato de $LogL$ ser bastante bem comportada, quando m é irrestrita. A estimação com m diferente de zero, às vezes, infla consideravelmente os erros-padrão dos demais parâmetros. Outras vezes, os leva a valores extremos e, com bastante frequência, impede ou dificulta a convergência das iterações. Segundo Greene (1993a), dado o estágio do estado-da-arte, é difícil desenhar qualquer conclusão sobre estas questões, pois não se conhece ainda como a restrição $m = 0$ afeta as estimativas de eficiência técnica.

Os erros-padrão aproximados das estimativas de máxima verossimilhança são calculados através da raiz quadrada dos elementos da diagonal da matriz de informação, oriunda da iteração final do algoritmo de otimização utilizado. Em geral, a matriz final de informação é uma boa aproximação da inversa do Hessiano da função logarítmica de verossimilhança, a menos que o algoritmo termine após umas poucas iterações (Judge et al. *apud* COELLI, 1995).

Pela função $LogL$ (2.19) definida para o modelo ALS com U_i 's seminormais (ver detalhes na seção 2.1), verifica-se que o referido modelo é parametrizado em termos de $I = \mathbf{s}_U / \mathbf{s}_V$ e de $\mathbf{s}^2 = \mathbf{s}_V^2 + \mathbf{s}_U^2$. Assim, a razão $\mathbf{s}_U^2 / \mathbf{s}^2$ é um indicador da influência do componente de ineficiência sobre a variância global. Segundo Greene (1993a), a variância de U_i é

$$[(p-2)/p]\mathbf{s}_U^2 \quad (3.40)$$

e, ao se decompor a variância total em seus dois componentes, a contribuição de U_i fica sendo igual a

$$\mathbf{g}^\bullet = \frac{\text{Variância de } U_i}{\text{Variância de } \mathbf{e}_i} = \frac{[(p-2)/p]\mathbf{s}_U^2}{[(p-2)/p]\mathbf{s}_U^2 + \mathbf{s}_V^2} \quad (3.41)$$

Para efeitos de estimação, é algumas vezes conveniente usar uma parametrização diferente. Battese & Corra (1977) substituem I por $\mathbf{g} = \mathbf{s}_U^2 / \mathbf{s}^2$, em que γ toma um valor entre zero e 1. Esta parametrização tem vantagens durante a estimação por MLE, porque o espaço do parâmetro γ pode ser pesquisado, a partir de um valor inicial apropriado, segundo um algoritmo de maximização. Sob a referida parametrização, a função logarítmica de verossimilhança é

$$LogL(\varphi, \mathbf{s}, \mathbf{g}) = \frac{n}{2} (\ln 2\mathbf{p} + \ln \mathbf{s}^2) + \sum_{i=1}^n \ln \Phi(-z_i) - \frac{1}{2\mathbf{s}^2} \sum_{i=1}^n \mathbf{e}_i^2 \quad (3.42)$$

onde: $z_i = \left[\frac{\mathbf{e}_i}{\mathbf{s}} \right] \sqrt{\frac{\mathbf{g}}{1-\mathbf{g}}}$ e

Φ é a função de densidade normal cumulativa.

3.1.3.3 Estimação pelo método dos mínimos quadrados ordinários corrigido

Os parâmetros do modelo de fronteira estocástica também podem ser estimados utilizando o segundo e o terceiro momentos centrais dos resíduos de OLS, \mathbf{m}_2 e \mathbf{m}_3 , respectivamente. Para o modelo seminormal, segundo Greene (1993a), as equações dos referidos momentos são

$$\mathbf{m}_2 = \left[\frac{p-2}{p} \right] \mathbf{s}_U^2 + \mathbf{s}_V^2 \quad (3.43)$$

$$e \quad m_3 = \sqrt{\frac{2}{p}} \left[1 - \left(\frac{4}{p} \right) \right] s_U^3 \quad (3.44)$$

A partir destas equações, s_U^2 e s_V^2 podem ser consistentemente estimados através dos resíduos de OLS, e , conforme mostra Waldman *apud* GREENE (1982), por

$$\hat{s}_U^2 = \left(\sqrt{\frac{p}{2}} \left(\frac{p}{p-4} \right) \left(\sum_{i=1}^n e_i^3 / n \right) \right)^{2/3} \quad (3.45)$$

$$e \quad \hat{s}_V^2 = \left[\left(\sum_{i=1}^n e_i^2 / n \right) - \frac{p-2}{p} s_U^2 \right] \quad (3.46)$$

donde se pode escrever

$$\hat{s}^2 = \hat{s}_U^2 + \hat{s}_V^2 \quad (3.47)$$

$$m_2 = \frac{\sum_{i=1}^n e_i^2}{n} \quad (3.48)$$

$$e \quad m_3 = \frac{\sum_{i=1}^n e_i^3}{n} \quad (3.49)$$

Desde que, por (3.40),

$$E[U_i] = (2/p)^{1/2} s_U \quad (3.50)$$

a correção do coeficiente de interseção, obtido por OLS, é

$$\hat{b}_{0ols} = \hat{b}_{0ols} + \hat{s}_U \sqrt{\frac{2}{p}} \quad (3.51)$$

Estes estimadores são consistentes, porém ineficientes em comparação com os estimadores de máxima verossimilhança. O grau da referida ineficiência é algo que ainda precisa ser determinado.

Os estimadores COLS do modelo de fronteira ALS seminormal, segundo a parametrização de Battese & Corra (1977), podem ser obtidos pelo manuseio direto das expressões deduzidas por Olson et al. (1980). De acordo com Coelli (1995), substituindo $l = \sqrt{g/(1-g)}$ nas expressões usadas pela equipe de Olson chega-se, depois dos necessários rearranjos, aos seguintes estimadores de s^2 , g e b_{0cols} :

$$\hat{s}^2 = m_2 + \frac{2}{p} \left[\sqrt{\frac{p}{2}} \cdot \frac{p}{(p-4)} m_3 \right]^{2/3} \quad (3.52)$$

$$\hat{g} = \hat{s}^{-2} \left[\sqrt{\frac{p}{2}} \cdot \frac{p}{(p-4)} m_3 \right]^{2/3} \quad (3.53)$$

$$e \quad \hat{b}_{\text{cols}} = \hat{b}_{\text{ols}} + \sqrt{\frac{2 \hat{g} \hat{s}^2}{p}} \quad (3.54)$$

onde m_2 e m_3 são os estimadores de \mathbf{m}_2 e de \mathbf{m}_3 , respectivamente.

Em seu estudo, Olson et al. (1980) alertam para o fato de que o estimador COLS pode não existir (de uma forma significativa) em algumas amostras, em virtude da ocorrência de falhas do Tipo I e do Tipo II. Estas falhas acontecem, respectivamente, quando \mathbf{I} se aproxima de zero ou do infinito. Isto é equivalente, de acordo com a última parametrização acima, a \mathbf{g} se aproximando de zero ou de 1. Assim, a falha do Tipo I verifica-se quando m_3 é positivo e a do Tipo II, quando m_2 é menor do que

$$\frac{(p-2)}{p} \left[\sqrt{\frac{p}{2}} \cdot \frac{p}{(p-4)} m_3 \right]^{2/3} \quad (3.55)$$

Um valor positivo para m_3 torna negativa a expressão inserida dentro dos colchetes na equação (3.52), impedindo o cálculo de uma estimativa para \mathbf{s}^2 . Já um erro do Tipo II, não impede que as estimativas sejam calculadas, mas proporciona uma estimativa de \mathbf{g} que excede a unidade. Isto só pode acontecer no caso do valor de \mathbf{s}_V^2 ou de \mathbf{s}_U^2 ser negativo, o que é teoricamente impossível. Algumas estimativas de \mathbf{g} , ligeiramente maiores do que 1, podem ser obtidas quando o verdadeiro valor de \mathbf{g} está bem próximo de 1. Nestes casos, considera-se a estimativa COLS de \mathbf{g} como sendo igual a 1.

O grupo de Olson, além dos resultados mencionados acima, utiliza métodos de Monte Carlo para comparar as propriedades dos estimadores COLS e MLE referentes a um modelo do tipo ALS. Eles verificam que, para os coeficientes de todos os regressores, exceto o coeficiente de interseção, existe muito pouca diferença entre COLS e MLE, e lembram que para esses coeficientes, COLS = OLS. Suas conclusões apontam o método COLS como preferível para amostras de tamanho menor do que 400, e para \mathbf{I} menor do

que 3,16. E que, mesmo para tamanhos de amostra e razões de variâncias maiores, a eficiência adicional do método MLE pode não ser suficiente para decidir enfrentar as dificuldades em computá-lo.

Mesmo assim, MLE tem sido usado em muito mais aplicações do que COLS. Segundo Coelli (1995), a falta de aceitação do estimador COLS na literatura aplicada é parcialmente explicada pela sua deficiência em oferecer uma estimativa de I , na presença de um erro do Tipo II. Este tipo de falha, sugere que I é infinito. O uso da parametrização de Battese & Corra (1977) não sofre deste problema, pois emprega

$$g = s_U^2 / (s_V^2 + s_U^2) \quad (3.56)$$

ao invés de $I = s_U / s_V$.

Assim, quando um erro do Tipo II se manifesta, g é considerado igual a 1, o que indica que os desvios até a fronteira são inteiramente devidos à ineficiência técnica.

Outra razão importante para o desinteresse no emprego do estimador COLS está ligada ao acesso aos erros-padrão das estimativas. Olson e seus associados desenvolvem um esboço de como esses erros assintóticos são derivados, mas não apresentam as expressões finais. Usando estes delineamentos, Coelli chega a resultados completos para a derivação dos erros-padrão de $b_{0\text{cols}}$, s^2 e g (para $b_{0\text{cols}}$, ver detalhes no final da seção 3.2.1.2).

3.1.3.4 Estimação pelo método da análise de correlação canônica corrigida⁷: proposição

A análise CCA tem sido muito utilizada para identificar e quantificar as associações entre dois conjuntos de variáveis, de uma maneira geral. No caso de UTD's que produzem múltiplos bens e serviços, estes dois conjuntos de variáveis são o vetor de produtos Y ⁸ e o vetor de insumos X da função de produção múltipla

⁷ Esta denominação, concebida na presente pesquisa, engloba os novos procedimentos metodológicos oferecidos para a estimação de funções de produção múltipla de fronteira estocástica, mediante a utilização da análise de correlação canônica.

⁸ Para fins de desenvolvimento teórico, o vetor Y é admitido com tamanho igual ou menor do que o vetor X , ou seja, $p \leq m$. Os procedimentos para o caso em que $p > m$ são encontrados em Vinod (1968).

$$A'Y_i = (B'X_i) + e_i \quad (3.57)$$

$$\text{ou} \quad (Y_i^C = A'Y_i) = (X_i^C = B'X_i) + e_i \quad (3.58)$$

onde: $A' = (a_1, \dots, a_p)$ e $B' = (b_1, \dots, b_m)$ são os vetores dos coeficientes⁹

canônicos a estimar e

$$Y_i^C \text{ e } X_i^C$$

são as combinações lineares de Y_i e de

X_i , respectivamente, que são conhecidas

como variáveis canônicas.

Os vetores aleatórios Y_i , X_i , Y_i^C e X_i^C tem expectâncias, covariâncias e variâncias designadas por

$$\begin{aligned} E(Y_i) &= m_i; \quad E(X_i) = m_{X_i}; \\ \text{Cov}(Y_i) &= \Sigma_{YY}; \quad \text{Cov}(X_i) = \Sigma_{XX}; \quad \text{Cov}(Y_i, X_i) = \Sigma_{YX} = \Sigma'_{XY}; \\ \text{Var}(Y_i^C) &= A' \text{Cov}(Y_i) A = A' \Sigma_{YY} A; \quad \text{Var}(X_i^C) = B' \text{Cov}(X_i) B = B' \Sigma_{XX} B \text{ e} \\ \text{Cov}(Y_i^C, X_i^C) &= A' \text{Cov}(Y_i, X_i) B = A' \Sigma_{YX} B \end{aligned} \quad (3.59)$$

A análise de correlação canônica (Johnson & Wichern, 1988) busca determinar os vetores de coeficientes A e B que maximizam a correlação entre as variáveis canônicas, ou seja,

$$\text{Corr}(Y_i^C, X_i^C) = r = \frac{A' \Sigma_{YX} B}{\sqrt{A' \Sigma_{YY} A} \sqrt{B' \Sigma_{XX} B}} \quad (3.60)$$

Isto é equivalente à determinação de A' e de B' que

$$\begin{aligned} &\text{maximiza a } \text{Cov}(Y_i^C, X_i^C) \\ &\text{sujeita a } \text{Var}(Y_i^C) = 1 = \text{Var}(X_i^C) \end{aligned} \quad (3.61)$$

Aplicando os multiplicadores de Lagrange, conforme mostra Ruparel (1998), este problema fica representado pela expressão

$$L = \text{Cov}(A'Y_i^C, B'X_i^C) + m_i (A' \Sigma_{YY} A - 1) + m_{X_i} (B' \Sigma_{XX} B - 1) \quad (3.62)$$

Através da manipulação algébrica das condições de primeira ordem, chega-se à equação

⁹ Estes coeficientes são os parâmetros referentes à tecnologia de produção. Quando se trabalha com as variáveis dos vetores X e Y logaritmadas, a tecnologia representada é Cobb-Douglas.

$$(\Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} - \mathbf{r}^2 I) B = 0 \quad (3.63)$$

onde \mathbf{r} denota a maior correlação canônica entre Y_i^C e X_i^C .

Para Vinod (1976), é bem conhecido que \mathbf{r} é a probabilidade limite do coeficiente de correlação múltipla, quando Y_i^C é regredido contra X_i^C . Como Y_i^C e X_i^C têm média zero e variância unitária para y_{qi} ($q = 1, \dots, p$) e x_{ki} ($k = 1, \dots, m$) estandardizados, \mathbf{r} é também a probabilidade limite dos coeficientes de regressão. Disto decorre que

$$Y_i^C = \mathbf{r} X_i^C + \mathbf{e}_i \quad (3.64)$$

onde se admite que as p variáveis de produção dependem das m variáveis referentes aos insumos, e não o contrário.

Através da decomposição da matriz

$$(\Sigma_{XX}^{-1} \Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX}) \quad (3.65)$$

obtem-se os auto-vetores, que constituem uma solução para B , e os auto-valores ou raízes características, que constituem uma solução para \mathbf{r}^2 . A solução para A é obtida resolvendo a equação

$$A = \frac{\Sigma_{YY}^{-1} \Sigma_{YX} B}{\mathbf{r}} \quad (3.66)$$

Quando as observações das variáveis integrantes dos vetores X e Y são padronizadas, as matrizes de covariância amostral Σ são matrizes de correlação R . Neste caso, os coeficientes de correlação canônica ao quadrado são as raízes características de

$$RC = R_{YY}^{-1/2} R_{XX} R_{XX}^{-1} R_{XY} R_{YY}^{-1/2} \quad (3.67)$$

Para implementar as soluções apontadas nas formulações (3.65 a 3.67), as matrizes Σ_{YY} ou R_{YY} , e Σ_{XX} ou R_{XX} , têm de ser invertidas. Quando as variáveis constituintes dos vetores Y e X são altamente correlacionadas entre si (isto é, não-ortogonais, quase colineares), as matrizes inversas ficam instáveis e muito sensíveis a pequenas mudanças nos dados básicos. Para resolver este problema de multicolinearidade, Vinod (1976) sugere a aplicação do modelo de regressão de cumeeira canônica (*canonical ridge regression model* - CRM).

O desenvolvimento seguido até aqui é bem conhecido na literatura especializada sobre o assunto. Neste ponto, é sugerida uma nova proposição concernente ao uso do método CCA para estimar funções de produção de fronteira com mais de uma variável dependente, tendo em vista a mensuração dos correspondentes índices individuais de eficiência técnica. A essência desta nova abordagem é a incorporação do vetor de resíduos compostos do modelo ALS, em que $\mathbf{e}_i = V_i + U_i$.

Segundo Vinod (1976), as variáveis y_{qi} e x_{ki} estão sujeitas aos erros $\mathbf{e}_{y_{qi}}$ e $\mathbf{e}_{x_{ki}}$, respectivamente, tais que

$$\begin{aligned} E(\mathbf{e}_{y_{qi}}) &= 0 \\ E(\mathbf{e}_{x_{ki}}) &= 0 \\ E(\mathbf{e}_{y_{qi}} \mathbf{e}_{y_{q'i}}) &= 0, \forall q' \neq q \\ E(\mathbf{e}_{x_{ki}} \mathbf{e}_{x_{k'i}}) &= 0, \forall k' \neq k \\ E(\mathbf{e}_{y_{qi}} y_{q'i}) &= E(\mathbf{e}_{x_{ki}} x_{k'i}) = 0 \end{aligned} \quad (3.68)$$

No referido estudo, Vinod usa o mesmo procedimento empregado em OLS para estimar o coeficiente canônico de interseção (que não está contido no vetor B). Por conseguinte, os resíduos canônicos provenientes da estimação dos vetores de coeficientes A e B de (3.57), por CCA, podem ser considerados oriundos de modelos com as mesmas propriedades estatísticas dos que geram resíduos de mínimos quadrados ordinários.

Desta forma, os parâmetros do modelo de fronteira estocástica seminormal de produção múltipla são estimados utilizando o segundo (m_2) e o terceiro (m_3) momentos centrais dos resíduos canônicos (\mathbf{e}_i), calculados através de (3.48) e de (3.49), respectivamente. Com base nestes valores, \mathbf{S}_U^2 e \mathbf{S}_V^2 são estimados consistentemente pelas equações (3.45) ou (3.56) e (3.46) ou (3.47), respectivamente. E, se a média dos U_i 's canônicos, similarmente ao caso dos U_i 's de mínimos quadrados, é determinada por

$$E[U_i] = (2 / \mathbf{p})^{1/2} \mathbf{S}_U$$

a correção para o coeficiente de interseção obtido por CCA, correspondentemente à correção do intercepto obtido por OLS, é calculada analogamente como em (3.51) e (3.54). Este procedimento é batizado, nesta pesquisa, de “análise de correlação canônica

corrigida” (CCCA). Assim, as versões canônicas de (3.51) e de (3.54) são, respectivamente,

$$\hat{\mathbf{b}}_{0\text{ccca}} = \hat{\mathbf{b}}_{0\text{cca}} + \hat{\mathbf{s}}_U \sqrt{\frac{2}{p}} \quad (3.69)$$

$$\text{e} \quad \hat{\mathbf{b}}_{0\text{ccca}} = \hat{\mathbf{b}}_{0\text{cca}} + \sqrt{\frac{2 \hat{\mathbf{g}} \hat{\mathbf{s}}^2}{p}} \quad (3.70)$$

Evidentemente, considera-se, também, que todas as demais propriedades inerentes aos estimadores COLS são aplicáveis aos estimadores CCCA.

3.1.3.5 Estimação da eficiência técnica

Quando da idealização do modelo ALS, em 1977, é equacionado o cálculo da eficiência técnica média para todas as observações, mas não a sua estimativa para cada UTD. A solução para o problema da decomposição é encontrada cinco anos mais tarde por Jondrow et al. (1982).

Eles derivam a distribuição condicional de U_i dado \mathbf{e}_i , a qual contém toda a informação que \mathbf{e}_i produz sobre U_i . Utilizando a média e a moda da referida distribuição como estimativas pontuais de U_i , eles deduzem as expressões para as distribuições seminormal e exponencial.

No caso seminormal, consideram

$$\begin{aligned} V_i &\sim N(0, \mathbf{s}_V^2) \text{ e} \\ U_i &\sim |N(0, \mathbf{s}_U^2)| \end{aligned} \quad (3.71)$$

Por simplicidade notacional abandonam o subscrito i e adotam as definições (já desenvolvidas no modelo ALS)

$$\begin{aligned} \mathbf{s}^2 &= \mathbf{s}_U^2 + \mathbf{s}_V^2 \\ U_{\bullet} &= -\mathbf{s}_U^2 \mathbf{e} / \mathbf{s}^2 \text{ e} \\ \mathbf{s}_{\bullet}^2 &= \mathbf{s}_U^2 \mathbf{s}_V^2 / \mathbf{s}^2 \end{aligned} \quad (3.72)$$

para provar que a distribuição condicionada de U dado \mathbf{e} é a distribuição de uma variável $N(\mathbf{m}_{\bullet}, \mathbf{s}_{\bullet}^2)$ truncada em zero.

Como estimativa pontual de U , eles usam tanto a média como a moda da sua distribuição condicionada, sendo a primeira dada por

$$E(U|\mathbf{e}) = \mathbf{m} + \mathbf{s} \cdot \frac{f(-\mathbf{m}/\mathbf{s})}{1 - \Phi(-\mathbf{m}/\mathbf{s})} \quad (3.73)$$

onde f e Φ representam as funções de densidade normal padronizada e normal cumulativa, respectivamente.

Para o estimador pontual de U (também seminormal), dado pela moda de sua distribuição condicionada, os referidos pesquisadores definem

$$\begin{aligned} M(U|\mathbf{e}) &= -\mathbf{e}(\mathbf{s}_U^2/\mathbf{s}^2) & \text{se } \mathbf{e} \leq 0 \\ &= 0 & \text{se } \mathbf{e} > 0 \end{aligned} \quad (3.74)$$

Os procedimentos, quando U segue uma distribuição exponencial, são idênticos ao caso da distribuição seminormal, com exceção de que U é suposto ter distribuição exponencial a um parâmetro, com densidade $f(U) = \exp(U/\mathbf{s}_U)/\mathbf{s}_U$.

Seis anos após, Battese & Coelli (1988) propõem uma generalização dos resultados acima para dados de painel e U_i 's com distribuição seminormal. Eles tomam por base a função de produção de fronteira

$$\ln y_{it} = \ln fpf(X_{it}, \phi) + \mathbf{e}_i \quad (3.75)$$

onde: $i=1, \dots, n$ e $t=1, \dots, T$ indexam as UTD's e os períodos de tempo, respectivamente;

y_{it} é o volume de produção da i -ésima UTD no t -ésimo período de tempo;

X_{it} é um vetor de funções dos insumos associados com a i -ésima UTD no t -ésimo período de tempo;

ϕ é o vetor de parâmetros a estimar e

$\mathbf{e}_i = V_i + U_i$, $U_i \leq 0$ tem as mesmas definições do modelo ALS.

Os referidos pesquisadores definem, então, a eficiência técnica de uma determinada UTD como igual à razão entre a sua produção média (em unidades originais),

a um dado nível de eficiência, e a correspondente produção média se U_i fosse igual a zero, ou seja,

$$ET_i = \frac{E(y_{it} | U_i, X_{it})}{E(y_{it} | U_i = 0, X_{it})} \quad (3.76)$$

Esta medida, necessariamente, tem valores entre zero e um. Se a eficiência técnica de uma dada firma é, por exemplo, igual a 0,85, significa que ela consegue produzir, em média, 85% da produção alcançada por uma firma inteiramente eficiente, cujo volume de insumos é o mesmo daquela.

Para a função de produção de fronteira expressada em unidades originais de produção, a eficiência técnica da i -ésima firma é definida por

$$ET_i = (\bar{X}_i \phi + U_i) (\bar{X}_i \phi)^{-1} \quad (3.77)$$

onde \bar{X}_i é o nível médio dos insumos para a i -ésima firma.

Por sua vez, a correspondente medida da eficiência técnica média de todas as firmas é

$$ET = 1 - \left\{ m + \frac{sf(-m/s)}{1 - \Phi(-m/s)} \right\} (\bar{X} \phi)^{-1} \quad (3.78)$$

No caso da função de produção de fronteira ser definida pelo logaritmo da produção, a produção da i -ésima firma no t -ésimo período é $\exp(y_{it})$, e a sua correspondente medida da eficiência técnica é

$$ET_i = \exp(U_i) \quad (3.79)$$

Esta medida é equivalente à razão entre a produção da i -ésima firma, em um dado período t ,

$$\exp(y_{it}) = \exp(X_{it} \phi + V_{it} + U_i) \quad (3.80)$$

e o correspondente volume de produção se U_i fosse igual a zero, qual seja,

$$\exp(X_{it} \phi + V_{it}) \quad (3.81)$$

Este índice de eficiência técnica não é dependente do nível de utilização dos insumos para uma dada firma, ao contrário do que ocorre com a avaliação por (3.77).

Com base nestas definições, Battese e Coelli concluem que a medida da eficiência técnica média de todas as firmas da indústria considerada, que corresponde a (3.79), é

$$ET = \left\{ \frac{1 - \Phi \left[\mathbf{s} - (\mathbf{m}/\mathbf{s}) \right]}{1 - \Phi \left(-\mathbf{m}/\mathbf{s} \right)} \right\} \exp \left(-\mathbf{m} + \frac{1}{2} \mathbf{s}^2 \right) \quad (3.82)$$

Para equacionar a estimação da eficiência técnica para uma dada firma, eles se valem da definição emitida em (3.76) e da percepção de que ela depende de inferências sobre a apropriada função de U_i para uma determinada amostra. Assim, definem a distribuição condicionada de U_i , dado $\mathbf{e}_i = \mathbf{e}_i^{10}$, como uma distribuição normal truncada em zero, com média igual a

$$\mathbf{m}_i^* \equiv (-\mathbf{s}^2 \bar{\mathbf{e}}_i + T^{-1} \mathbf{m} \mathbf{s}_V^2) (\mathbf{s}^2 + T^{-1} \mathbf{s}_V^2)^{-1} \quad (3.83)$$

e variância
$$\mathbf{s}_i^2 \equiv \mathbf{s}^2 \mathbf{s}_V^2 (\mathbf{s}_V^2 + T \mathbf{s}^2)^{-1} \quad (3.84)$$

onde
$$\bar{\mathbf{e}}_i = T^{-1} \sum_{t=1}^T \mathbf{e}_{it} \quad (3.85)$$

Por seu turno, as expectativas condicionadas de U_i e de $\exp(U_i)$, dado $\mathbf{e}_i = \mathbf{e}_i$, são determinadas, respectivamente, por

$$E(U_i | \mathbf{e}_i = \mathbf{e}_i) = \mathbf{m}_i^* + \mathbf{s}_i^2 \left\{ \mathbf{f}(-\mathbf{m}_i^* / \mathbf{s}_i^2) [1 - \Phi(-\mathbf{m}_i^* / \mathbf{s}_i^2)]^{-1} \right\} \quad (3.86)$$

e
$$E[\exp(U_i) | \mathbf{e}_i = \mathbf{e}_i] = \left\{ \frac{1 - \Phi \left[\mathbf{s}_i^2 - (\mathbf{m}_i^* / \mathbf{s}_i^2) \right]}{1 - \Phi \left(-\mathbf{m}_i^* / \mathbf{s}_i^2 \right)} \right\} \exp \left(-\mathbf{m}_i^* + \frac{1}{2} \mathbf{s}_i^2 \right) \quad (3.87)$$

Nesta formulação, os autores evidenciam que a média \mathbf{m}_i^* , definida por (3.83), é de ordem 1 e converge, quando $T \rightarrow \infty$, até o limite

$$\lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T \mathbf{e}_{it} = U_i$$

que é o valor de U_i para a i -ésima UTD. E que a variância \mathbf{s}_i^2 , definida por (3.84), é de ordem T^{-1} e converge para zero, quando $T \rightarrow \infty$.

¹⁰ Neste caso, \mathbf{e}_i é o resíduo da i -ésima UTD obtido através da função de produção de fronteira estocástica estimada por COLS ou MLE e, a partir da validação dos resultados da presente pesquisa, por CCCA.

Assim, uma vez conhecidos os parâmetros da função de produção de fronteira definida em unidades originais de produção, o estimador de Battese e Coelli da eficiência técnica da i -ésima UTD, de acordo com (3.77), é

$$\hat{U}_i \equiv M_i^\bullet + \hat{\mathbf{s}}_\bullet \left\{ \mathbf{f}(-M_i^\bullet / \hat{\mathbf{s}}_\bullet) [1 - \Phi(-M_i^\bullet / \hat{\mathbf{s}}_\bullet)] \right\}^{-1} \quad (3.88)$$

onde M_i^\bullet é a variável aleatória amostral correspondente à média \mathbf{m}^\bullet , definida em (3.83), sendo igual a

$$M_i^\bullet \equiv (-\mathbf{s}^2 \bar{\mathbf{e}}_i + T^{-1} \mathbf{m} \mathbf{s}_V^2) (\mathbf{s}^2 + T^{-1} \mathbf{s}_V^2)^{-1} \quad (3.89)$$

Para o caso em que a função de produção de fronteira é definida em termos do logaritmo da produção, os referidos pesquisadores apresentam o seu estimador da eficiência técnica da i -ésima UTD, conforme definido por (3.79), como sendo

$$\hat{ET}_i = \left[\exp(\hat{U}_i) | \hat{\mathbf{e}}_i = \hat{\mathbf{e}}_i \right] = \left\{ \frac{1 - \Phi \left[\hat{\mathbf{s}}_\bullet - (M_i^\bullet / \hat{\mathbf{s}}_\bullet) \right]}{1 - \Phi(-M_i^\bullet / \hat{\mathbf{s}}_\bullet)} \right\} \exp(-M_i^\bullet + \frac{1}{2} \hat{\mathbf{s}}_\bullet^2) \quad (3.90)$$

que é obtido pela substituição de \mathbf{m}^\bullet em (3.87), por M_i^\bullet definido em (3.89).

Sobre a robustez do estimador proposto, Battesi e Coelli demonstram que ele é não-tendencioso, pois $E(\hat{U}_i) = E(U_i)$, e é consistente quando $T \rightarrow \infty$, pois (3.89) converge em probabilidade para $\text{plim}_{T \rightarrow \infty} -\bar{\mathbf{e}}_i = U_i$, e a variável aleatória $\mathbf{f}(-M_i^\bullet / \hat{\mathbf{s}}_\bullet) [1 - \Phi(-M_i^\bullet / \hat{\mathbf{s}}_\bullet)]^{-1}$, converge em probabilidade para zero. Os referidos pesquisadores também mostram que os resultados de Jondrow et al. (1982), para o caso seminormal e dados de *cross-section*, são obtidos quando $\mathbf{m} = 0$ e $T = 1$ nas equações (3.83), (3.84) e (3.86). Segundo Battesi e Coelli, Jondrow e sua equipe somente deixam de explicitar a expressão da expectância condicionada, $\exp(U_i) | \mathbf{e}_i = \mathbf{e}_i$, dada na equação (3.87).

3.2 MODELOS DE ANÁLISE

Nesta seção, definem-se e justificam-se as especificações gerais seguidas no tocante ao número de equações do modelo, ao método de estimação, à forma funcional, à estocasticidade da fronteira, à distribuição dos resíduos representativos da ineficiência, à natureza dos dados amostrais e à verificação de hipóteses econométricas. Por conveniência

organizacional, estas especificações são apresentadas sob dois títulos, reservando-se ao segundo apenas as questões inerentes à verificação de hipóteses econométricas.

Também, apresentam-se as especificações singulares necessárias à estimação de cada modelo, especialmente quanto à definição das variáveis integrantes, à explicitação da fronteira estocástica e ao equacionamento dos correspondentes índices de eficiência técnica para cada UTD. As hipóteses econométricas específicas fazem-se necessárias no caso da fronteira de custo e da fronteira de produção múltipla. Elas tratam, basicamente, de verificar a existência de efetiva dualidade no primeiro caso. No segundo, visam avaliar a significância do ajuste canônico e o cumprimento das condições de primeira e de segunda ordens.

A diretriz básica que norteia as especificações adotadas é, estritamente, assegurar condições uniformes que garantam comparabilidade aceitável aos resultados obtidos com os diferentes modelos estimados, principalmente no confronto com o novo modelo. O trabalho assumiria proporções excessivamente extensas se considerasse vários tipos funcionais, vários métodos de estimação, vários tipos de distribuição dos resíduos representativos da ineficiência, e assim por diante.

3.2.1 Especificações gerais

3.2.1.1 Delineamento básico

Nesta pesquisa utiliza-se, exclusivamente, funções a uma equação, pois o uso de modelos multiequacionais na análise de eficiência está “na sua infância”, havendo ainda muito para ser feito (Lovell, 1993). A abordagem a mais de uma equação abriga facilmente o caso de tecnologias de produção com múltiplos produtos e consiste, basicamente, na estimação simultânea de funções de fronteira de custo (de lucro ou de receita) e de equações de demanda por fator a elas associadas, seguindo o caminho aberto por Christensen & Greene (1976). De acordo com William H. Greene¹¹, grande parte da pesquisa econométrica recente, referente à mensuração de eficiência, está concentrada na especificação e estimação desse tipo de modelo. Seu ajustamento, contudo, defronta-se

com formidáveis dificuldades na modelagem da relação entre as perturbações bi-caudais das equações de demanda por fator e a parte residual não-negativa (ineficiência alocativa) presente na função primária. Estas dificuldades caracterizam o problema conhecido na literatura como *“Greene problem”*, em homenagem a quem primeiramente o percebeu (Greene, 1980). Até hoje o referido problema ainda não está completamente resolvido.

Para Chizmar & Zak (1984), o sistema de equações simultâneas com uma equação para cada produto, em que cada um deles é produzido em processos separados, é o meio mais correto para estimar funções de produção com múltiplos produtos. Todavia, eles reconhecem que ao se colocar cada produto como uma função dos insumos (inclusive de um termo residual), exclui-se as implicações mais interessantes advindas da produção de mais de um produto em um único processo produtivo. Além disso, na estimação de tais modelos, passa-se a enfrentar dificuldades equivalentes às apontadas no final do parágrafo anterior. Assim, pelas razões expostas neste parágrafo e no anterior, a presente pesquisa utiliza somente modelos econométricos uniequacionais.

A estimação destes modelos, para fins terminais de mensuração da eficiência técnica, não é feita por MLE, porque este método ainda não está desenvolvido para estimar diretamente funções de fronteira com mais de uma variável dependente. Por isso, opta-se pelo emprego de estimadores COLS para todos os modelos antigos. Para o modelo novo, adota-se uma extensão do método COLS à análise CCA denominada, de forma pioneira nesta pesquisa, de “análise de correlação canônica corrigida (CCCCA)”.

No que tange à escolha do tipo de função, inicialmente decidiu-se adotar duas formas funcionais básicas. A conhecida função Cobb-Douglas padrão, que é linear nos logaritmos das variáveis, homogênea, homotética e tem elasticidade de substituição unitária. E a função transcendental logarítmica, mais conhecida como translog, que é quadrática nos logaritmos das variáveis, não-homogênea, não-homotética e tem elasticidade de substituição não-unitária. A função translog é bastante flexível para aproximar tecnologias de produção arbitrárias em termos de possibilidades de substituição.

¹¹ Informações verbais prestadas ao autor da presente pesquisa durante a sua permanência no Departamento de Economia da Universidade de Nova York, em meados de 1997.

Não obstante, deixou-se de utilizá-la porque a sua principal vantagem reside na estimação conjunta com sistemas de equações de demanda por fator.

Com relação ao emprego de funções que representem unicamente a estrutura da tecnologia de produção, Bauer (1990) argumenta que geralmente são utilizadas fronteiras estocásticas porque elas permitem a consideração dos ruídos estatísticos resultantes de eventos que fogem do controle da firma, tais como a sorte e o tempo, além de erros de especificação e de variáveis omitidas (não correlacionadas com os regressores incluídos). Para ele, a obtenção de estimativas de eficiência é mais concernente com uma fronteira estocástica do que com uma fronteira determinística, na qual as estimativas provêm diretamente dos resíduos da estimação. Além disso, segundo Greene (1993b), a fronteira determinística vem sendo abandonada por motivos práticos.

Por sua vez, o uso de fronteiras estocásticas requer a adoção de uma determinada distribuição para o componente residual indicativo de ineficiência técnica, ou seja, o termo U_i da função de produção de fronteira estocástica definida em (3.33). Segundo Vitaliano & Toren (1994), na falta de um maior direcionamento oriundo da teoria econômica, os econométricos têm sido guiados por conveniência e intuição. A distribuição seminormal é a mais largamente utilizada, mas também têm sido empregadas em menor escala a truncada normal e a exponencial. A distribuição gama é uma alternativa atraente (Greene, 1990), mas seu uso geral carece de programas para tratamento computacional. A literatura especializada neste assunto é bem recente e ainda não está completamente desenvolvida. Até o momento, as propriedades amostrais finitas dos estimadores COLS foram investigadas somente para o caso de funções de produção de fronteira estocástica seminormal (Coelli, 1995). Como o presente trabalho vale-se estritamente de estimadores COLS ou assemelhados, assume-se em todo o seu desenvolvimento que os U_i 's têm distribuição seminormal.

Quando se dispõe de um painel de dados, não há necessidade de definir uma distribuição específica para o componente residual unicaudal. Se a eficiência não varia ao longo do tempo, tem-se no painel um instrumento útil para estimar a eficiência técnica de cada UTD considerada. Porém, esta é uma suposição demasiado forte, particularmente quando o número de observações da série temporal aumenta (Bauer, 1990). O

relaxamento da hipótese de invariabilidade impõe a especificação de um modelo sobre como a eficiência varia com o transcorrer do tempo. Isto não é muito complicado, contudo torna a tarefa de estimação bem mais árdua. Na presente pesquisa, além destas implicações, o desconhecimento das propriedades amostrais finitas dos estimadores COLS para painel de dados (Coelli, 1995), torna proibitivo o seu uso. Consequentemente, a pesquisa tanto no seu desenvolvimento metodológico quanto na análise empírica contempla, exclusivamente, dados de *cross-section*.

3.2.1.2 Verificação de hipóteses

Em resumo, o desenvolvimento metodológico da pesquisa é embasado em:

- modelos econométricos uniequacionais;
- tecnologias de produção multi-produto tipo Cobb-Douglas;
- fronteiras estocásticas;
- resíduos seminormais;
- dados de *cross-section*;
- estimação por COLS e CCCA e
- mensuração exclusiva da eficiência técnica relativa.

De acordo com esta seqüência de especificações gerais, a primeira verificação que se impõe é a da existência efetiva de uma fronteira estocástica, ou da hipótese nula de que não existe ineficiência técnica. Isto pode ser levado a efeito através do teste de hipóteses

$$\begin{aligned} H_0: \mathbf{s}_U^2 &= 0 \\ H_1: \mathbf{s}_U^2 &> 0 \end{aligned} \quad (3.91)$$

ou, equivalentemente,

$$\begin{aligned} H_0: \mathbf{I} \text{ ou } \mathbf{g} &= 0 \\ H_1: \mathbf{I} \text{ ou } \mathbf{g} &> 0 \end{aligned} \quad (3.92)$$

Desde a idealização do modelo ALS, estas hipóteses vêm sendo testadas, em quase todas as aplicações, por intermédio da estatística devida a Wald (Coelli, 1995), que é dada pela razão das estimativas de \mathbf{s}_U^2 , \mathbf{I} ou de \mathbf{g} com o respectivo erro-padrão estimado. Quando são utilizadas estimativas obtidas por MLE, uma alternativa ao teste de Wald é o teste da razão de verossimilhança. Este teste requer a estimação do modelo sob

cada uma das hipóteses, nula ($\text{Log}L_0$) e alternativa ($\text{Log}L_1$), sendo sua estatística calculada por

$$LR = -2(\text{Log}L_0 - \text{Log}L_1) \quad (3.93)$$

Pesquisando sobre os estimadores e os testes de hipóteses para funções de fronteira estocástica, através de experimentos de Monte Carlo, Coelli (1995) conclui que o teste de Wald e o teste LR bilateral têm abrangência incorreta. No caso do teste LR , o problema é evitado pelo uso de um teste LR unicaudal que tem alcance correto e potência superior à do teste de Wald.

Contudo, a realização do teste LR unilateral requer a estimação do modelo de fronteira estocástica por MLE. Schmidt & Lin (1984), pesquisando a construção de um teste com as estimativas de OLS, definem uma estatística com o multiplicador de Lagrange. Porém, a tentativa não prospera porque, sob a hipótese $H_0: I = 0$, o vetor das primeiras derivadas parciais da função $\text{Log}L$ é um vetor zero e a matriz de informação é singular. Ao chegar a esta conclusão eles se voltam para outras direções.

Os referidos pesquisadores observam, então, que a existência ou não de uma fronteira estocástica, com base na estimação por OLS, envolve essencialmente o teste de que $e_i = V_i + U_i$ seja normalmente distribuído. Apesar de reconhecer a validade dos testes de normalidade existentes, apegam-se à constatação de que a soma de uma distribuição normal com uma seminormal deve sofrer de falta de assimetria. Propõem, então, o emprego do teste de significância do terceiro momento dos resíduos de OLS, ou seja, da inclinação

$$\sqrt{b_1} = \frac{m_3}{m_2^{2/3}} \quad (3.94)$$

Este teste tem a vantagem de não requerer o cálculo de estimativas de máxima verossimilhança da fronteira estocástica, tem dimensão correta, tem potência quase equivalente a do teste LR unicaudal (Coelli, 1995) e conta com a distribuição de $\sqrt{b_1}$ discutida e tabulada em alguns trabalhos, como o de D'Agostino & Pearson (1973). Por isso, o referido teste é usado em todos os modelos da pesquisa, para testar a existência de cada tipo de fronteira.

Uma vez comprovada a hipótese de não-normalidade dos resíduos compostos de mínimos quadrados, e definida a distribuição seminormal para o componente unicaudal (conforme a seção anterior), resta ainda a tarefa de testar a hipótese de que a especificação adotada para U_i seja verdadeira.

A confirmação de que os resíduos indicativos da eficiência técnica sejam reconhecidamente seminormais, não tem pautado a grande maioria das aplicações empíricas realizadas. Os pesquisadores, via de regra, têm se limitado a assumir U_i 's seminormais sem qualquer verificação. Nesta pesquisa, deixa-se de testar outras possibilidades como a gama/normal e a gama/exponencial, devido à inexistência de *software* apropriado. As referidas distribuições ainda estão muito pouco desenvolvidas para a aplicação em questão.

Na formulação do modelo ALS, seus idealizadores efetuam a referida verificação mediante a comparação dos valores maximizados das funções logarítmicas de verossimilhança ($\text{Log}L$). Deste modo, decidem que a distribuição seminormal indica um ajuste melhor do que a exponencial, em um dos dois exemplos empíricos apresentados.

Este procedimento, ou mesmo a utilização dos valores $\text{Log}L$ em testes LR , de acordo com Greene (1991), é inapropriado. O referido autor adverte que o espaço dos parâmetros das distribuições sob comparação e, portanto, das respectivas funções de verossimilhança, não são relacionados.

Na presente pesquisa, a seminormalidade dos U_i 's é avaliada para cada um dos l modelos de análise, através de observação visual dos histogramas e de testes de aderência tipo qui-quadrado (Costa Neto, 1977). Nestes testes, é examinada a hipótese nula de que os resíduos unicaudais seguem distribuições normal truncada e exponencial, utilizando a estatística dada pela fórmula

$$\mathbf{c}_{j(\text{calc})}^2 = \sum_{cl=1}^{CL} \frac{(fo_{cl} - fe_{cl})^2}{fe_{cl}} \quad (3.95)$$

onde: $\mathbf{c}_{j(\text{calc})}^2$, $j = 1, \dots, l$, é comparado com o valor crítico, $\mathbf{c}_{j(CL-2)}^2$, com $CL-2$

graus de liberdade;

CL é o número de classes ou intervalos;

fo_{cl} são as frequências observadas nas classes e
 fe_{cl} são as frequências esperadas nas classes. Elas são U_i 's gerados com parâmetros extraídos dos dados amostrais, segundo distribuições teóricas do tipo normal truncada e exponencial.

A derradeira tarefa da fase de verificação econométrica geral é medir a significância individual dos coeficientes de regressão dos modelos estimados. Para testar a significância dos coeficientes de interseção estimados por COLS, emprega-se estatísticas t calculadas com as fórmulas dos erros-padrão assintóticos derivadas por Coelli (1995). Este autor, dando continuidade aos esforços de Olson et al. (1980), deduz aproximações para a variância de $\hat{\mathbf{b}}_{0cols}$, mediante o uso da expansão da série de Taylor, truncada após os termos de primeira ordem, e obtém o seguinte resultado:

$$\text{Var}(\hat{\mathbf{b}}_{0cols}) = \left[\frac{\mathbf{1} \hat{\mathbf{b}}_{0cols}}{\hat{\mathbf{b}}_{0cols}} \right]^2 \text{Var}(\hat{\mathbf{b}}_{0ols}) + \left[\frac{\mathbf{1} \hat{\mathbf{b}}_{0cols}}{\mathbf{1} m_3} \right]^2 \text{Var}(m_3) \quad (3.96)$$

onde: $\left[\frac{\mathbf{1} \hat{\mathbf{b}}_{0cols}}{\hat{\mathbf{b}}_{0ols}} \right] = 1;$

$$\left[\frac{\mathbf{1} \hat{\mathbf{b}}_{0cols}}{\mathbf{1} m_3} \right] = \frac{\mathbf{p}}{(\mathbf{p} - 4)} \cdot \frac{1}{3} \left[\sqrt{\frac{\mathbf{p}}{2}} \cdot \frac{\mathbf{p}}{(\mathbf{p} - 4)} m_3 \right]^{-2/3} e$$

$$\text{Var}(m_3) = \frac{1}{n} (m_6 - m_3^2 - 6m_2m_4 + 9m_2^3).$$

Coelli (1995) também apresenta os seguintes estimadores dos momentos centrais dos resíduos de OLS, em termos de \mathbf{s}^2 e de \mathbf{g} :

$$\begin{aligned} m_2 &= \hat{\mathbf{s}}^2 \left[(1 - \mathbf{g}) + \mathbf{g} \frac{(\mathbf{p} - 2)}{\mathbf{p}} \right]; \\ m_3 &= \hat{\mathbf{s}}^3 \left[\sqrt{\frac{2}{\mathbf{p}}} \left(1 - \frac{4}{\mathbf{p}} \right) \mathbf{g}^{3/2} \right]; \\ m_4 &= \hat{\mathbf{s}}^4 \left[3(1 - \mathbf{g})^2 + 6 \frac{(\mathbf{p} - 2)}{\mathbf{p}} \mathbf{g}(1 - \mathbf{g}) + \left(3 - \frac{4}{\mathbf{p}} - \frac{12}{\mathbf{p}^2} \right) \mathbf{g}^2 \right] e \end{aligned} \quad (3.97)$$

$$m_6 = \hat{S}^6 \left[15 (1-g)^2 + 45 \frac{(p-2)}{p} (1-g)^2 g + 15 \left(3 - \frac{4}{p} - \frac{12}{p^2} \right) (1-g) g^2 \dots \right. \\ \left. \dots + \left(15 - \frac{6}{p} - \frac{100}{p^2} - \frac{40}{p^3} \right) g^3 \right]$$

Para testar a significância dos coeficientes dos regressores utilizam-se as estatísticas t calculadas com erros-padrão oriundos das estimações por OLS¹², de conformidade com a orientação dada por Coelli (1999). Como para esses coeficientes existe bem pouca diferença entre OLS e MLE (Olson et al., 1980), pode-se empregar também os erros-padrão provindos de estimações por MLE. Este último procedimento, porém, não pode ser aplicado na estimação das fronteiras de produção múltipla e de distância-produto. No primeiro caso, porque inexistente algoritmo de maximização da função $LogL$ estocástica com mais de uma variável dependente. E, no segundo, porque a estimação econométrica impõe transformações na variável dependente, as quais só podem ser incorporadas com a mudança do *software* de estimação por MLE.

3.2.2 Especificações particulares

Nesta seção, discrimina-se cada modelo de análise envolvido adotando-se para todos eles a forma funcional Cobb-Douglas, cuja escolha é fundamentada na seção 3.2.1.1. Também, são definidos os testes econométricos específicos que são necessários em decorrência de características singulares.

A função de produção tipo Cobb-Douglas é bastante conhecida dado o seu uso consagrado na estimação empírica das relações de insumo-produto, carecendo de qualquer apresentação.

3.2.2.1 Fronteira de produção agregada

Na construção da fronteira de produção agregada, pressupõe-se que existam n UTD's, utilizando um vetor de insumos $X \in \Re_{++}^m$ para produzir um produto (bem e/ou serviço) escalar agregado $y \in \Re_{++}$, com um vetor de preços finais

¹² No caso da estimação por CCCA, lança-se mão da correspondência dos resíduos de OLS com os de CCA, conforme constatado na seção 3.1.3.4, para empregar os erros-padrão provenientes da estimação por CCA.

$PF = (pf_1, \dots, pf_p) \in \Re_{++}^p$ e tecnologia Cobb-Douglas dada pela seguinte função de produção agregada de fronteira estocástica

$$\ln y_i = \mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + (V_i + U_i) \quad (3.98)$$

onde: $\mathbf{b}_{0\text{cols}}$ é o coeficiente de interseção corrigido a estimar por (3.51) ou (3.54);

\mathbf{b}_k é o k -ésimo parâmetro a estimar por OLS e

$U_i \leq 0$ é o indicador de ineficiência técnica da i -ésima UTD.

De acordo com (3.36), a fronteira estocástica de produção agregada correspondente à função (3.98) é

$$\left(\mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + V_i \right) \quad (3.99)$$

Esta fronteira é determinada pela fronteira determinística definida conforme (3.35),

$\left(\mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} \right)$, e por acontecimentos externos que estão fora do controle do produtor (V_i). Enquadrando-a na forma geral (3.37), define-se a eficiência técnica de cada UTD, originada pela fronteira de produção agregada (3.99), como sendo

$$\ln ET_i = \ln y_i - \left(\mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + V_i \right) = U_i \quad \therefore \quad ET_i = e^{U_i} \quad (3.100)$$

À semelhança de (2.21), agrega-se a variável representativa do conjunto da produção através da utilização de

$$y_i^{\text{pai}} = \sum_{q=1}^p y_{qi} \quad (3.101)$$

como um índice da variável dependente y_i , e inclui-se $p-1$ variáveis independentes adicionais, dadas por

$$y_{qi}^{\text{pai}} = \frac{y_{qi}}{\sum_{q=1}^p y_{qi}} \quad (3.102)$$

Desta forma, as expressões (3.98) e (3.100), quando a agregação é feita através do índice acima, são, respectivamente,

$$\ln y_i^{\text{pai}} = \mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \sum_{q=1}^{p-1} \mathbf{a}_q^{\text{pai}} \ln y_{qi}^{\text{pai}} + (V_i + U_i) \quad (3.103)$$

e

$$\ln ET_i = \ln y_i^{\text{pai}} - \left(\mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \sum_{q=1}^{p-1} \mathbf{a}_q^{\text{pai}} \ln y_{qi}^{\text{pai}} + V_i \right) = U_i \quad \therefore ET_i = e^{U_i} \quad (3.104)$$

Quando o produto é agregado por intermédio dos coeficientes de correlação canônica, o processo de produção é especificado na forma implícita como $FI(Y_i, X_i) = 0$, com $Y_i = (y_1, \dots, y_p)$ produtos e $X_i = (x_1, \dots, x_m)$ insumos. Para fins de estimação, restringe-se a função para satisfazer a forma funcional $g(Y_i) = f(X_i)$, cujas funções representam a curva de transformação e a função de produção múltipla, respectivamente, da i -ésima UTD que produz p bens Y_i . Usando m tipos de insumos X_i , ela busca minimizar o custo total de produção, transformando os bens em produtos com o objetivo de maximizar a sua receita total (Rao, 1969).

Para modelar esta formulação, define-se, seguindo (2.25), a função Cobb-Douglas na forma estendida, como sendo

$$\sum_{q=1}^p \mathbf{a}_q \ln y_{qi} = \mathbf{b}_{0\text{cca}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \mathbf{e}_i \quad (3.105)$$

Após estimar esta equação por CCA, de acordo com a seção 3.1.3.4, agrega-se o produto por meio da operação

$$\sum_{q=1}^p \hat{\mathbf{a}}_q \ln y_{qi} \quad (3.106)$$

onde $\hat{\mathbf{a}}_q$ é o q -ésimo coeficiente canônico estimado. Então, fazendo

$$\ln y_i^{\text{pac}} = \sum_{q=1}^p \hat{\mathbf{a}}_q \ln y_{qi} = \ln y_i \quad (3.107)$$

retorna-se à equação (3.98) para estimá-la por COLS.

Com o intuito de padronizar nomenclaturas e siglas, os dois modelos abrangidos nesta seção ficam denominados fronteira de produção agregada por índice (PAI) e fronteira de produção agregada por correlação canônica (PAC).

3.2.2.2 Fronteira de custo

Na especificação da fronteira de custo (CP), considera-se que existam n UTD's, cada qual com um vetor de custos igual a $CP = (w_1x_1, \dots, w_mx_m) = (cp_1, \dots, cp_m) \in \mathfrak{R}_{++}^m$ referente à utilização de um conjunto de insumos $X \in \mathfrak{R}_{++}^m$, a preços $W = (w_1, \dots, w_m) \in \mathfrak{R}_{++}^m$, para produzir um conjunto de produtos e/ou serviços $Y \in \mathfrak{R}_{++}^p$, com tecnologia Cobb-Douglas determinada pela função de custo de fronteira estocástica

$$\ln(cp_i/w_{mi}) = b_{0cols} + \frac{1}{k} \sum_{q=1}^p a_q \ln y_{qi} + \sum_{k=1}^{m-1} \frac{h_k}{k} \ln(w_{ki}/w_{mi}) + \frac{1}{k} (V_i + U_i) \quad (3.108)$$

onde: $cp_i = \sum_{k=1}^m cp_k$;

b_{0cols} é o coeficiente de interseção corrigido a estimar por (3.109);

a_q e $\frac{h_k}{k}$ são o q -ésimo e o k -ésimo parâmetros a estimar por OLS, respectivamente;

$k = \sum_{k=1}^m h_k$ é a medida dos retornos de escala;

$U_i \geq 0$ é o indicador de ineficiência técnica da i -ésima UTD e

$\frac{1}{k}$ é o fator utilizado para eliminar o efeito de economias ou deseconomias de escala.

Segundo Schmidt & Lovell (1979), a eliminação dos efeitos de escala requer a divisão do segundo momento dos resíduos de mínimos quadrados, por k^2 , e do terceiro momento, por k^3 ; e a multiplicação de s_U^2 e de s_V^2 , por k^2 , conforme as expressões (2.36) e (2.37), respectivamente. E, para a correção do coeficiente de interseção obtido por OLS¹³, os referidos autores propõem o emprego de (2.39), ou, equivalentemente,

¹³ Em (3.109), não se faz uso da parametrização de Battese & Corra (1977), para conservar a formulação original de Schmidt & Lovell (1979). É fácil verificar que $s_U \sqrt{2/p} = \sqrt{2 g s^2 / p}$.

$$\hat{\mathbf{b}}_{0\text{cols}} = \hat{\mathbf{b}}_{0\text{ols}} - \frac{1}{\hat{\mathbf{k}}} \hat{\mathbf{s}}_U \sqrt{\frac{2}{p}} \quad (3.109)$$

que difere de (3.51) somente quanto ao fator $\frac{1}{\hat{\mathbf{k}}}$ e à troca de sinal.

No entanto, as providências para neutralizar os efeitos de escala sobre a fronteira de custo são dispensáveis na presente pesquisa. Por um lado, porque não produzem nenhuma alteração nos parâmetros \mathbf{s}_U^2 e \mathbf{s}_V^2 . E, por outro, porque sua aplicação na correção do coeficiente de interseção modifica a magnitude dos índices de eficiência técnica, porém, sem mudar o posto relativo de cada UTD na amostra.

A equação (3.108) segue as concepções de Schmidt & Lovell (1979) e de Kumbhakar (1989), explicitadas na seção 2.3. Com base na adaptação de (3.36) para o caso de uma função de custo, especifica-se

$$\left[\mathbf{b}_{0\text{cols}} + \frac{1}{\mathbf{k}} \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} + \sum_{k=1}^{m-1} \frac{\mathbf{h}_k}{\mathbf{k}} \ln (w_{ki}/w_{mi}) + \frac{1}{\mathbf{k}} (V_i) \right] \quad (3.110)$$

como sendo a fronteira estocástica de custo. Esta fronteira é determinada pela estrutura da tecnologia de custo adaptada de (3.35), $\left[\mathbf{b}_{0\text{cols}} + \frac{1}{\mathbf{k}} \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} + \sum_{k=1}^{m-1} \frac{\mathbf{h}_k}{\mathbf{k}} \ln (w_{ki}/w_{mi}) \right]$, e por acontecimentos externos que estão fora do controle do produtor (V_i).

A eficiência técnica individual para cada firma, definida de forma geral em (3.37) para uma fronteira de produção, no caso de avaliação via fronteira de custo (3.110), é dada por

$$\ln ET_i = \ln(cp_i/w_{mi}) - \left[\mathbf{b}_{0\text{cols}} + \frac{1}{\mathbf{k}} \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} + \sum_{k=1}^{m-1} \frac{\mathbf{h}_k}{\mathbf{k}} \ln(w_{ki}/w_{mi}) + \frac{1}{\mathbf{k}} V_i \right] = U_i \quad \therefore \quad ET_i = e^{U_i} \quad (3.111)$$

É importante observar que na estimação de ET_i deve-se trocar o sinal de \bar{e}_i na formulação original (3.89), a qual é desenvolvida para uma fronteira de produção e não de custo.

A interpretação da função de custo de fronteira deve levar em conta o enfoque dual da estimação. Ou seja, deve-se evitar que um produtor considerado tecnicamente eficiente, através de mensuração via função de produção, seja considerado alocativamente ineficiente, se avaliado mediante a correspondente função de custo.

Por isso, no modelo de custo de fronteira estocástica, considera-se uma bem comportada tecnologia de produção com possibilidade de eficiência técnica apenas, ou seja, sob a suposição de exata minimização de custos. Para testar esta hipótese, basta verificar se a participação de cada insumo no custo é igual para todas as observações da amostra. Se a minimização de custos é exata, então,

$$\frac{w_{ki}x_{ki}}{w_{li}x_{li}}, \quad k = 2, \dots, m \quad (3.112)$$

deve ser igual para todas as UTD's (Schmidt & Lin, 1984).

Este teste depende só da suposição de idêntica e homotética tecnologia para cada observação. É um teste com probabilidade zero de erro. Por uma questão prática, aceita-se a hipótese nula em casos em que a participação dos insumos no custo não seja exatamente igual, mas esteja bem próxima da igualdade.

Outras condições que a função de custo deve atender, para corresponder a uma bem comportada estrutura de produção, são homogeneidade linear, monotonicidade e convexidade nos preços dos insumos.

A primeira condição implica em que a função de custo deve ser homogênea de grau 1 nos preços dos insumos. Esta exigência está contemplada na função de custo de fronteira (3.108), pois seus criadores incluem tal requisito na sua especificação (Schmidt & Lovell, 1979).

Para que as duas outras condições se verifiquem, é suficiente que as participações no custo sejam positivas e o Hessiano, limitado à função de custo, seja negativo. O atendimento destas requisições é providenciado na geração dos dados amostrais (ver seção 3.3) por meio de funções monotônicas com elasticidade constante de substituição e com elasticidade constante de transformação, de conformidade com as recomendações de Kumbhakar (1989) e de Färe et al. (1994).

Assim, dado que as condições de homogeneidade linear, de monotonicidade e de convexidade são satisfeitas por construção, não há necessidade de especificar testes para sua verificação econométrica.

3.2.2.3 Fronteira de produção inversa

Para discriminar a fronteira de produção inversa (PI), supõe-se que n produtores podem usar um único tipo de insumo $x \in \mathfrak{R}_{++}$ para produzir um conjunto de produtos e/ou serviços $Y \in \mathfrak{R}_{++}^p$, com tecnologia Cobb-Douglas especificada pela função de produção inversa de fronteira estocástica

$$\ln x_i = \mathbf{b}_{0\text{cols}} + \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} + (V_i + U_i) \quad (3.113)$$

onde: $\mathbf{b}_{0\text{cols}}$ é o coeficiente de interseção corrigido a estimar por (3.109), sem o

fator $\frac{1}{\bar{k}}$;

\mathbf{a}_q é o q -ésimo parâmetro a estimar por OLS e

$U_i \geq 0$ é o indicador de ineficiência técnica da i -ésima UTD.

A fronteira estocástica de produção inversa derivada da função (3.113) é demarcada pela fronteira determinística, $\left(\mathbf{b}_{0\text{cols}} + \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} \right)$, definida segundo (3.35) e por acontecimentos externos que fogem ao controle do produtor (V_i). Considerando a versão de (3.36) para o caso de uma função de produção inversa, tem-se a seguinte fronteira estocástica:

$$\left(\mathbf{b}_{0\text{cols}} + \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} + V_i \right) \quad (3.114)$$

Assim, a expressão geral da eficiência técnica de cada unidade produtora representada em (3.37) para uma fronteira de produção, quando a mensuração se dá por meio da fronteira de produção inversa (3.114), é traduzida por

$$\ln ET_i = \ln x_i - \left(\mathbf{b}_{0\text{cols}} + \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} + V_i \right) = U_i \quad \therefore \quad ET_i = e^{U_i} \quad (3.115)$$

A função de produção inversa é suposta ser não-decrescente, quase côncava e contínua para baixo (Diewert, 1974). Ela define a fronteira de possibilidades de produção, dado o nível de x . Pode ser vista como uma função de custo, onde o custo (em termos do insumo x) é expressado como uma função dos níveis de produto Y . Também, pode ser interpretada como a função de necessidades de insumo, a qual dá o montante mínimo de insumo x requerido para produzir o vetor de produtos Y .

O uso da função de produção inversa requer uma suposição de complementariedade (proporções fixadas) entre todos os principais insumos necessários à produção. Na seção 2.4, observa-se que Gathon & Perelman (1992) e Kumbhakar & Hjalmarsson (1995) fazem uso dos insumos com participações de mais de 90 e 80% nos custos totais de produção, respectivamente, para justificar a escolha de um único insumo. Eles alegam, também, que, nessas condições, as possibilidades de substituição entre os insumos escolhidos e os demais, são bastante limitadas.

Quando a restrição de complementariedade não é atendida, pode-se empregar a variante mencionada por Lovell (1995) e agregar os insumos através de um índice. Para isto, utiliza-se o mesmo procedimento definido para agregar o produto, visto na seção 3.2.2.1. Assim, a variável dependente representativa do conjunto dos insumos x_i , seguindo (3.108), é dada por

$$x_i^{pi} = \sum_{k=1}^m x_{ki} / w_{mi} \quad (3.116)$$

e inclui-se $m-1$ variáveis independentes adicionais, dadas por

$$x_{ki}^{pi} = \frac{x_{ki}}{x_i^{pi}} \quad (3.117)$$

Desta forma, (3.113) fica sendo igual a

$$\ln x_i^{pi} = b_{0cols} + \sum_{q=1}^p a_q \ln y_{qi} + \sum_{k=1}^{m-1} b_k^{pi} \ln x_{ki}^{pi} + (V_i + U_i) \quad (3.118)$$

e (3.115) igual a

$$\ln ET_i = \ln x_i^{pi} - \left(b_{0cols} + \sum_{q=1}^p a_q \ln y_{qi} + \sum_{k=1}^{m-1} b_k^{pi} \ln x_{ki}^{pi} + V_i \right) = U_i \therefore ET_i = e^{U_i} \quad (3.119)$$

3.2.2.4 Fronteira de distância-produto

Pela função de distância-produto, almeja-se o maior incremento proporcional possível no vetor observado de produtos, considerando que o vetor expandido deve ainda ser um elemento do conjunto de produção original $P(X)$, ou seja, repetindo (3.13),

$$D_{\text{out}}(X, Y) = \min\{\mathbf{d}: (Y / \mathbf{d}) \in P(X)\}$$

Esta função é não-decrescente, positiva, linearmente homogênea e convexa em Y , e decrescente em X . Por definição, e conforme já enunciado em (3.14) e (3.15), tem-se, respectivamente, que

$$D_{\text{out}}(X, Y) \leq 1 \Leftrightarrow Y \in P(X)$$

e
$$D_{\text{out}}(X, Y) = 1 \Leftrightarrow Y \in \text{Isoq } P(X)$$

onde $\text{Isoq } P(X)$ é a fronteira do conjunto de produção.

A função de distância-produto carrega todas as propriedades da tecnologia-mãe, para a qual oferece uma representação funcional que descreve completamente a referida tecnologia. Ao contrário da função de produção, ela prontamente modela tecnologias com múltiplos produtos.

A referida função é dual para a função receita e permite chegar-se aos preços-sombra mediante seu uso direto (via lema da dualidade de Sheppard), requerendo somente dados sobre quantidades, tanto de insumos quanto de produtos. Ainda que se possa derivar preços de produtos (custo marginal) usando uma função de custo, este enfoque exige dados sobre preços de todos os insumos.

Para modelar a fronteira de distância-produto, assume-se que n UTD's empregam um vetor de insumos $X \in \mathfrak{R}_{++}^m$ para produzir um vetor de produtos $Y \in \mathfrak{R}_{++}^p$, de acordo com a tecnologia de produção dada pela correspondência-produto, a qual mapeia cada $X \in \mathfrak{R}_{++}^m$ em um conjunto de produção $P(X) \subseteq \mathfrak{R}_{++}^p$. Esta correspondência satisfaz os principais axiomas da produção (vistos na seção 3.1.1) e oferece estrutura e regularidade suficientes ao uso da teoria da dualidade.

Para fins de estimação, recorre-se às propriedades de homogeneidade e de fechamento como fazem Lovell et al. (1990), e define-se a função Cobb-Douglas de distância-produto de fronteira estocástica¹⁴

$$\ln(y_{pi}^{-1}) = \mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \sum_{q=1}^{p-1} \mathbf{a}_q^{\text{dp}} \ln(y_{qi}/y_{pi}) + (V_i + U_i) \quad (3.120)$$

onde: $\mathbf{b}_{0\text{cols}}$ é o coeficiente de interseção corrigido a estimar por (3.51) ou (3.54), considerando os fatores de correção com sinais negativos;

\mathbf{b}_k e \mathbf{a}_q^{dp} são o k -ésimo e o q -ésimo parâmetros a estimar por OLS, respectivamente e

$U_i \geq 0$ é o indicador de ineficiência técnica da i -ésima UTD.

De acordo com (3.36), a fronteira estocástica de distância-produto (DP), originada de (3.120), é

$$\left[\mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \sum_{q=1}^{p-1} \mathbf{a}_q^{\text{dp}} \ln(y_{qi}/y_{pi}) + V_i \right] \quad (3.121)$$

Ela é determinada, de conformidade com (3.35), pela estrutura da tecnologia de produção,

$\left[\mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \sum_{q=1}^{p-1} \mathbf{a}_q^{\text{dp}} \ln(y_{qi}/y_{pi}) \right]$, e por fatores externos alheios à intervenção do produtor (V_i).

Empregando (3.37) particulariza-se o conceito geral da eficiência técnica de firmas individuais através da fronteira de distância-produto, a qual fica explicitada por

$$\ln ET_i = \ln(y_{pi}^{-1}) - \left[\mathbf{b}_{0\text{cols}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \sum_{q=1}^{p-1} \mathbf{a}_q^{\text{dp}} \ln(y_{qi}/y_{pi}) + V_i \right] = U_i \quad \therefore ET_i = e^{U_i} \quad (3.122)$$

3.2.2.5 Fronteira de produção múltipla: proposição

Na modelagem da fronteira de produção múltipla (PM), emprega-se a suposição de que n produtores usam um vetor de insumos $X \in \mathfrak{R}_{++}^m$ para produzir um conjunto de

¹⁴ Este procedimento também é recomendado por Grosskopf (1998), em resposta ao e-mail a ela

bens e/ou serviços $Y \in \mathfrak{R}_{++}^p$, com tecnologia representada pela função de produção múltipla de fronteira estocástica, na forma Cobb-Douglas estendida

$$\sum_{q=1}^p \mathbf{a}_q \ln y_{qi} = \mathbf{b}_{0\text{cca}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + (V_i + U_i) \quad (3.123)$$

onde: $\mathbf{b}_{0\text{cca}}$ é o coeficiente de interseção corrigido a estimar por (3.69) ou (3.70);

\mathbf{a}_q e \mathbf{b}_k são o q -ésimo e o k -ésimo parâmetros a estimar por CCA e

$U_i \leq 0$ é o indicador de ineficiência técnica da i -ésima UTD.

Assim, por (3.36), a quantidade

$$\left(\mathbf{b}_{0\text{cca}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + V_i \right) \quad (3.124)$$

é a fronteira estocástica de produção múltipla, que é determinada pela fronteira determinística, $\left(\mathbf{b}_{0\text{cca}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} \right)$, definida em (3.35), e por eventos externos não controláveis pela UTD (V_i). Aquela fronteira, quando processada de acordo com (3.37), evidencia a eficiência técnica de firmas individuais. Para isto, faz-se

$$\ln ET_i = \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} - \left(\mathbf{b}_{0\text{cca}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + V_i \right) = U_i \therefore ET_i = e^{U_i} \quad (3.125)$$

Por seu turno, a função de produção múltipla

$$\sum_{q=1}^p \mathbf{a}_q \ln y_{qi} = \mathbf{b}_{0\text{cca}} + \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} + \mathbf{e}_i \quad (3.126)$$

é estimada por CCA de conformidade com a seção 3.1.3.4, através da escolha dos coeficientes canônicos \mathbf{a}_q 's e \mathbf{b}_k 's que maximizam a correlação entre as variáveis canônicas

$$Y_i^C = \sum_{q=1}^p \mathbf{a}_q \ln y_{qi} \quad (3.127)$$

$$\text{e} \quad X_i^C = \sum_{k=1}^m \mathbf{b}_k \ln x_{ki} \quad (3.128)$$

Quando se trabalha com as variáveis de produção e de insumos padronizadas (Vinod, 1976), o coeficiente de interseção é igual a

$$b_{0cca} = \sum_{q=1}^p \overline{\ln y_{qi}} - \sum_{k=1}^m \left(\frac{b_k}{s_{\ln x_{ki}}} \right) \overline{\ln x_{ki}} \quad (3.129)$$

onde: $\overline{\ln x_{ki}}$ e $\overline{\ln y_{qi}}$ são as médias aritméticas dos logaritmos das variáveis x_{ki} e y_{qi} , respectivamente e $s_{\ln x_{ki}}$ é o desvio -padrão do k -ésimo insumo considerado.

A avaliação da significância estatística global do ajustamento canônico é, usualmente, realizada através de testes de hipótese que empregam as distribuições de qui-quadrado e F . De acordo com Johnson & Wichern (1988), quando $\sum_{YX} = 0$, as variáveis canônicas Y_i^C e X_i^C têm covariância $A' \Sigma_{YX} B = 0$, para todos os vetores A e B . Consequentemente, todas as correlações canônicas são iguais a zero e não existe nenhuma razão para se efetuar a análise CCA. Assim, para n e $n-(p+m)$ grandes, ele orienta que se pode rejeitar

$$H_0: \Sigma_{YX} = 0 \quad (r_1 = r_2 = \dots = r_p = 0) \quad (3.130)$$

a um determinado nível de significância (n.s), desde que

$$c_{calc}^2 = -(n-1 - \frac{1}{2}(p+m+1)) \ln \prod_{q=1}^p (1 - \hat{r}_q^2) > c_{pm}^2 (n.s) \quad (3.131)$$

onde $c_{pm}^2 (n.s)$ é o valor crítico do centésimo percentil da distribuição de qui-quadrado com $(p.m)$ graus de liberdade.

Se a hipótese nula (3.130) é rejeitada, deve-se examinar a significância das correlações canônicas individuais. Como as correlações canônicas são ordenadas decrescentemente, no teste seguinte supõe-se que a primeira correlação canônica é diferente de zero, e que as remanescentes $p-1$ correlações são iguais a zero, ou seja,

$$H_0: r_1 \neq 0 \quad (r_2 = \dots = r_p = 0) \quad (3.132)$$

Esta hipótese é testada a um certo nível de significância pelo qui-quadrado com $(p-1)(m-1)$ graus de liberdade. Se ela é aceita, admite-se que as duas primeiras correlações canônicas sejam diferentes de zero e que as $p-2$ correlações remanescentes

sejam iguais a zero, e assim por diante. Este tipo de teste é conhecido como “teste de qui-quadrado de Bartlett”.

Outro teste, também muito utilizado, emprega a distribuição F e é denominado “teste F de Rao”. Para Vinod (1968), a hipótese nula (3.130) pode ser investigada através da estatística

$$F_{\text{calc}} = \frac{\left(1 - \left(\prod_{q=1}^p (1 - \hat{\mathbf{r}}_q^2) \right)^{1/sr} \right)}{\left(\prod_{q=1}^p (1 - \hat{\mathbf{r}}_q^2) \right)^{1/sr}} \left\| \left(\frac{mr \cdot sr + 2 \cdot vr}{p \cdot m} \right) \right\| \quad (3.133)$$

onde: F tem $(p \cdot m)$ graus de liberdade no numerador e $(mr \cdot sr + 2 \cdot vr)$ graus de liberdade no denominador;

$$mr = \left[n - 1 - \frac{1}{2}(p + m + 1) \right];$$

$$sr = \left(\frac{p^2 \cdot m^2 - 4}{p^2 + m^2 - 5} \right)^{1/2} \text{ e}$$

$$vr = \left[(-1) \left(\frac{p \cdot m - 2}{4} \right) \right].$$

As demais hipóteses, enunciadas para o teste de qui-quadrado de Bartlett, também podem ser testadas através do teste F de Rao, bastando, para isso, apenas considerar as diferentes situações em que \mathbf{r} é testado. As tabelas-padrão das distribuições de \mathbf{c}^2 e de F podem ser utilizadas em todos os testes de significância do ajustamento canônico.

Se a significância global do ajustamento canônico é estatisticamente satisfatória, é de utilidade mensurar a intensidade com que as variáveis canônicas representam seus respectivos conjuntos. Isto é feito através do cálculo do coeficiente de determinação (R^2) que mede a proporção da variância amostral total que é explicada pelas variáveis canônicas. Segundo Johnson & Wichern (1988), o referido coeficiente é computado através da fórmula

$$R^2_{y_{qi} | \hat{y}_i^c} = \frac{1}{p} \sum_{q=1}^p r_{\hat{y}_i^c, y_{qi}}^2 \quad (3.134)$$

para medir o poder de explicação da variável \hat{Y}_i^C , e por

$$R^2_{x_{ki}|\hat{X}_i^C} = \frac{1}{p} \sum_{k=1}^m r^2_{\hat{X}_i^C, x_{ki}} \quad (3.135)$$

para medir o poder de explicação da variável \hat{X}_i^C . Em cada caso, r expressa o coeficiente de correlação amostral entre cada variável canônica e as respectivas variáveis de produção ou de insumos. No caso de as observações estarem padronizadas, $r_{\hat{Y}_i^C, y_{qi}}$ é o q -ésimo elemento da matriz de correlação amostral de \hat{Y}_i^C com y_{qi} , ou seja,

$$R_{\hat{Y}_i^C, y_{qi}} = A' R_{YY} \quad (3.136)$$

e $r^2_{\hat{X}_i^C, x_{ki}}$ é o k -ésimo elemento da matriz de correlação amostral de \hat{X}_i^C com x_{ki} , ou seja,

$$R_{\hat{X}_i^C, x_{ki}} = B' R_{XX} \quad (3.137)$$

Um problema que persiste no método CCA é a sua deficiência quanto aos testes de significância estatística. Isto resulta da inexistência de desenvolvimento teórico para os erros-padrão dos parâmetros estimados. Gyimah-Brempong & Gyapong (1991), propõem a estimação dos erros-padrão dos coeficientes canônicos por intermédio do método Delta. No entanto, em resposta a uma consulta sobre o referido método, Gyimah-Brempong (1999) aconselha o uso do método de reamostragem conhecido como *bootstrap*.

Ruggiero (1999), opinando sobre a referida questão, recomenda o uso do método de reamostragem chamado *jackknife*. Através da abordagem desenvolvida por Judge et al. (1988), verifica-se que não há predominância importante de um método sobre o outro, quando aplicados a estimadores de mínimos quadrados.

Por razão de simplicidade, adota-se na pesquisa o método *jackknife*. Sua técnica consiste em obter $n-1$ estimativas dos erros-padrão, através de $n-1$ regressões canônicas. Em cada uma destas regressões, elimina-se da amostra uma linha de observações e repõe-se a linha excluída na regressão anterior. Assim, a variabilidade das $n-1$ estimativas é usada como a estimativa da variabilidade do estimador original (Judge et al., 1988).

Na presente pesquisa, a aplicação da análise CCA, a uma função de produção múltipla de fronteira estocástica, é inédita. Por isso, além da avaliação da qualidade do

ajuste canônico, deve-se examinar as implicações econômicas relacionadas com as condições de primeira e de segunda ordens. Estes aspectos são discutidos a seguir, de acordo com os desenvolvimentos encontrados em Vinod (1976) e Gyimah-Brempong & Gyapong (1991).

A elasticidade (EMg) e a produtividade marginais (PMg) de cada produto com relação a cada insumo são, respectivamente,

$$EMg(y_q, x_k) = \frac{\frac{\partial \ln y_q}{\partial \ln x_k}}{\frac{\partial \ln y_q}{\partial \ln x_k}} = \frac{b_k}{a_q} \quad (3.138)$$

$$e \quad PMg(y_{qi}, x_{ki}) = (y_{qi} / x_{ki}) EMg(y_{qi}, x_{ki}) = (y_{qi} / x_{ki}) \left(\frac{b_k}{a_q} \right) \quad (3.139)$$

De acordo com a teoria econômica da produção, estas grandezas marginais devem ser positivas na região de interesse para a pesquisa. Estas condições não ocorrem quando as variáveis de produção e de insumos são altamente correlacionadas entre si. Neste caso, deve-se empregar o modelo de regressão CRM.

O comportamento da função de possibilidades de produção múltipla, que é o lugar geométrico da função implícita

$$FI(Y_i, X_i) = \sum_{q=1}^p a_q \ln y_{qi} - b_{0cca} + \sum_{k=1}^m b_k \ln x_{ki} - e_i = 0 \quad (3.140)$$

para valores fixados de x_{ki} , é observado através do cálculo da taxa marginal de transformação entre dois produtos, dada por

$$TMgTP(y_{qi}, y_{q'i}) = \frac{\frac{\partial \ln y_{qi}}{\partial \ln y_{q'i}}}{\frac{\partial \ln y_{qi}}{\partial \ln y_{q'i}}} = - \left[\left(\frac{a_{q'}}{y_{q'i}} \right) / \left(\frac{a_q}{y_{qi}} \right) \right] \quad (3.141)$$

onde $q \neq q'$.

Com o volume de insumos fixado em certo nível, é intuitivamente óbvio que maior quantidade de um produto deve resultar em menor quantidade do outro. Para isso, é preciso que a curva de transformação do produto seja declinante. Analogamente, é necessário que a taxa marginal de substituição entre dois insumos

$$TMgSI(x_{ki}, x_{k'i}) = \frac{\frac{\partial \ln x_{ki}}{\partial \ln x_{k'i}}}{\frac{\partial \ln x_{ki}}{\partial \ln x_{k'i}}} = - \left[\left(\frac{b_{k'}}{x_{k'i}} \right) / \left(\frac{b_k}{x_{ki}} \right) \right] \quad (3.142)$$

seja também negativa, o que significa que as isoquantas, isto é os lugares geométricos de x_{ki} e de $x_{k'i}$ para valores fixados de y_{qi} e de $y_{q'i}$, devem ser decrescentes (notar que $k \neq k'$ e $q \neq q'$).

Por sua vez, a condição de convexidade das isoquantas segue um bem conhecido requisito da teoria microeconômica, pelo qual

$$\text{Conv}(x_{li}, \dots, x_{mi}) = \begin{bmatrix} 2 \left(\frac{\mathbb{I}FI}{\mathbb{I}x_{li}} \right) \dots \left(\frac{\mathbb{I}FI}{\mathbb{I}x_{mi}} \right) \left(\frac{\mathbb{I}^2 FI}{\mathbb{I}x_{li} \dots \mathbb{I}x_{mi}} \right) \\ - \left(\frac{\mathbb{I}FI}{\mathbb{I}x_{li}} \right)^2 \left(\frac{\mathbb{I}^2 FI}{\mathbb{I}x_{li}^2} \right) - \dots - \left(\frac{\mathbb{I}FI}{\mathbb{I}x_{mi}} \right)^2 \left(\frac{\mathbb{I}^2 FI}{\mathbb{I}x_{mi}^2} \right) \end{bmatrix} < 0 \quad (3.143)$$

Já a curva de transformação entre dois produtos é côncava para a origem se

$$\text{Conv}(y_{li}, \dots, y_{pi}) = \begin{bmatrix} 2 \left(\frac{\mathbb{I}FI}{\mathbb{I}y_{li}} \right) \dots \left(\frac{\mathbb{I}FI}{\mathbb{I}y_{pi}} \right) \left(\frac{\mathbb{I}^2 FI}{\mathbb{I}y_{li} \dots \mathbb{I}y_{pi}} \right) \\ - \left(\frac{\mathbb{I}FI}{\mathbb{I}y_{li}} \right)^2 \left(\frac{\mathbb{I}^2 FI}{\mathbb{I}y_{li}^2} \right) - \dots - \left(\frac{\mathbb{I}FI}{\mathbb{I}y_{pi}} \right)^2 \left(\frac{\mathbb{I}^2 FI}{\mathbb{I}y_{pi}^2} \right) \end{bmatrix} < 0 \quad (3.144)$$

Esta última condição pressupõe preços constantes para os produtos, o que leva à curvas de iso-receita com formato de linhas retas. No entanto, a ocorrência de preços variáveis é mais realista e resulta em curvas de iso-receita convexas à origem. Quando isto acontece, as condições de segunda ordem equacionadas acima são positivas devido à presença de economias de escala.

3.3 DADOS AMOSTRAIS

Na revisão da literatura, apresentada no capítulo 2, verifica-se que é muito raro encontrar, em um mesmo estudo, a utilização de mais de um modelo de fronteira. Isto se deve ao fato de que é muito difícil conseguir um determinado conjunto de UTD's, cuja matriz dos dados observados comporte a estimação de todos os modelos de análise, sem deixar de atender aos requisitos estatísticos, econométricos e de teoria econômica, conforme detalhados para cada modelo no capítulo 3. Por esta razão, opta-se por empregar uma base de dados simulada, de conformidade com a metodologia definida nas próximas duas seções.

3.3.1 Metodologia básica

A tecnologia de referência (*benchmark*), usada para construir a base de dados da pesquisa, é a tecnologia idealizada por Färe et al. (1994). Ela é caracterizada por elasticidade constante de transformação nos produtos (CET) e elasticidade constante de substituição nos insumos do tipo Cobb-Douglas (CD)¹⁵. Todos os possíveis vetores de insumos e de produtos, ou seja, a representação gráfica desta tecnologia é expressada pela função $FI(Y_i, X_i) = g(Y_i) - f(X_i) = 0$, na qual

$$g(Y_i) = \left(\frac{1}{2} y_{1i}^2 + \frac{1}{2} y_{2i}^2 \right)^{\frac{1}{2}} \quad (3.145)$$

$$\text{e} \quad f(X_i) = \left(x_{1i}^{1/2} x_{2i}^{1/2} \right)^k, \quad k > 0 \quad (3.146)$$

considerando dois produtos e dois insumos. A extensão para mais de dois produtos e para mais de dois insumos é direta.

Segundo Färe e sua equipe, a forma CET-CD é consistente com a convexidade requerida para $L(Y)$ e $P(X)$ ¹⁶, em que variando o parâmetro k permite-se a ocorrência de retornos variáveis de escala.

¹⁵ A escolha desta tecnologia deve-se à sua consagração na literatura como tecnologia padrão. Para ela, converge a grande maioria das funções de produção, quando submetidas a determinadas restrições.

¹⁶ Kumbhakar (1989) alerta para o problema da tecnologia de produção Cobb-Douglas não satisfazer as condições usuais de convexidade para maximização de lucros, no caso de funções com múltiplos produtos e múltiplos insumos. Recomenda, então, que isto seja evitado pela adoção de funções com elasticidade constante de substituição e com elasticidade constante de transformação.

Com base em (3.145) e (3.146), eles constroem uma amostra de 20 observações, dispostas em 4 grupos de 5 observações cada. Para o primeiro grupo de pequenos produtores, adotam

$$g(Y_i) = 25, \quad f(X_i) = 25 \quad \text{e} \quad k = 0,898.$$

Para os próximos dois grupos, de médios produtores, definem

$$g(Y_i) = f(X_i) = 50 \text{ ou } 75 \quad \text{e} \quad k = 1,000.$$

E, para o último grupo, de grandes produtores,

$$g(Y_i) = f(X_i) = 100 \quad \text{e} \quad k = 0,927.$$

Assim, a tecnologia simulada é integrada por regiões que apresentam rendimentos de escala crescente, constante e decrescente.

Os correspondentes dados de custos são obtidos da função de custo cp_i , derivada de acordo com a tecnologia acima representada. O lema de Shephard é empregado para gerar os dados sobre os preços dos insumos (w_1 e w_2). Deste modo, a referida função e seus termos componentes são os seguintes:

$$cp_i = 2 \left(\frac{1}{2} y_{1i}^2 + \frac{1}{2} y_{2i}^2 \right)^{1/2 k} (w_1 w_2)^{1/2} \quad (3.147)$$

$$\begin{aligned} w_{1i} &= \frac{cp_i}{2 x_{1i}} \\ w_{2i} &= \frac{cp_i}{2 x_{2i}} \end{aligned} \quad (3.148)$$

3.3.2 Geração e seleção das amostras aleatórias

A amostra de 20 observações, produzida por Färe et al. (1994), de acordo com a metodologia apresentada na seção anterior, é demasiado reduzida para os propósitos da presente pesquisa. Por isso, é preciso expandir o número de observações, mantendo a mesma tecnologia, os mesmos parâmetros e os mesmos intervalos daquela amostra. Nestas condições, a tarefa de simulação é configurada exclusivamente para produzir e selecionar amostras aleatórias, com base nos seguintes procedimentos:

I) Gerar dois sub-conjuntos de dados da variável y_i , com 115 observações cada, segundo distribuições uniformes. Os limites inferior e superior de cada sub-conjunto são os da metodologia básica.

II) Empilhar os dois sub-conjuntos em um único vetor de tamanho 230¹⁷, objetivando atender o requisito de não-normalidade dos resíduos obtidos por OLS e por CCA.

III) Gerar um conjunto de dados com 230 observações - distribuídas uniformemente dentro dos limites indicados na metodologia básica - para:

- as funções $g(Y_i)$ e $f(X_i)$, conjuntamente tituladas de h ;
- o parâmetro de variação de escala k e
- as variáveis x_{li} e cp_i .

IV) Criar 230 observações sobre y_2 e sobre x_2 por meio do processamento das equações abaixo, deduzidas de (3.145) e de (3.146), respectivamente,

$$y_{2i} = \sqrt{\frac{h^2 - 0,5 y_{li}^2}{0,5}} \quad (3.149)$$

$$e \quad x_{2i} = \exp \left[\frac{(\ln h - k \cdot 0,5 \ln x_{li})}{k \cdot 0,5} \right] \quad (3.150)$$

V) Reproduzir os correspondentes dados de preços dos insumos, por (3.148), e descartar as linhas de observações em que y_2 tenha apresentado valores iguais a zero.

VI) Estimar por OLS os modelos especificados na seção 3.2.2, exceto o de produção múltipla que é estimado por CCA. Calcular para cada modelo o valor de $\sqrt{b_1}$, por (3.94), e do coeficiente de determinação (R^2)¹⁸. Para isto, utilizar a matriz de dados resultante da implementação dos itens I a V.

VII) Repetir todos os procedimentos anteriores para gerar uma outra amostra, e assim por diante. Este processo é encerrado no momento em que a soma¹⁹ de

¹⁷ O tamanho inicial da matriz de dados artificiais é de 230 observações (linhas). Este número, de acordo com ensaios preliminares, garante matrizes finais com pelo menos $n \geq 150$, após o descarte previsto no item V desta seção. É adotado este dimensionamento porque na grande maioria das pesquisas consultadas são empregadas amostras com tamanho $n > 100$.

¹⁸ $\sqrt{b_1}$ e R^2 são indicadores da importância da amostra quanto à qualidade estatística das fronteiras estimadas e dos ajustamentos efetuados, respectivamente. O cálculo de R^2 em regressões por OLS é bastante conhecido, sendo seu valor um *output* obrigatório do *software* de estimação. No caso das regressões por CCA, ele é obtido pela média aritmética das expressões (3.134) e (3.135).

¹⁹ Para efeito desta soma, os valores positivos de $\sqrt{b_1}$, no caso das funções de custo e de produção inversa, são multiplicados por (-1).

$\sqrt{b_1}$, referente a todos os modelos de análise, por um bom período de tempo de processamento, não logre ultrapassar os valores máximos registrados em amostras anteriores. Para evitar a superlotação da área de armazenagem, guarda-se apenas as amostras em que as referidas somas fiquem gravitando perto de valores máximos.

VIII) Selecionar como a amostra única da pesquisa, dentre as amostras armazenadas, a que acusar a maior soma de $\sqrt{b_1}$.

3.4 ANÁLISE COMPARATIVA

Em consonância com a motivação central da pesquisa, a avaliação dos seus resultados não precisa ir além do cotejamento dos índices originados pelos diferentes modelos, em termos de correlação ordinal e de variabilidade.

Estudos similares, comparando as técnicas DEA e CCA ou utilizando o chamado CCA híbrido (CCA/DEA), têm sido validados mediante o emprego de testes não-paramétricos sobre a qualidade estatística da correlação ordinal e sobre a consistência das diferentes classificações. O ordenamento dos índices por postos, também é feito na presente pesquisa, todavia, a abordagem não-paramétrica é seguida devido à não-normalidade dos resíduos indicativos de eficiência técnica. Para levar a cabo este tipo de análise, usa-se o coeficiente de correlação (bivariada) ordinal de Spearman, que é dado por

$$SP_{j,j'} = SP_{j',j} = 1 - \frac{6 \sum_{j,j'=1, i=1}^{l,n} (RK_{ji} - RK_{j'i})^2}{n^3 - n}, \quad j \neq j', \quad j, j' = 1, \dots, l \quad (3.151)$$

onde: RK_{ji} é o posto atribuído à i -ésima UTD com índice de eficiência técnica estimado pelo j -ésimo modelo e

$RK_{j'i}$ é o posto atribuído à i -ésima UTD com índice de eficiência técnica medido pelo j' -ésimo modelo.

Nesta formulação, as linhas da matriz de índices são integradas pelas UTD's, e as colunas, pelos diferentes modelos. Assim, o ordenamento é realizado segundo cada modelo, ou seja, por colunas.

Para testar a hipótese nula de que a correlação entre os postos dos índices (dois a dois) seja igual a zero, compara-se o valor calculado de $SP_{j,j'}$ com o respectivo valor crítico tabelado. Este teste pode ser feito através da variável

$$t_{j,j'} = t_{j',j} = \frac{SP_{j,j'}\sqrt{n-2}}{\sqrt{1-SP_{j,j'}^2}}, \quad j \neq j', \quad j, j' = 1, \dots, l \quad (3.152)$$

cujas distribuições são aproximadamente iguais à distribuição t com $n-2$ graus de liberdade.

O coeficiente de correlação ordinal de Spearman é apenas um indicador do grau de associação linear entre os postos dos índices de eficiência das UTD's, segundo cada modelo (dois a dois). Para avaliar até que ponto os escores de eficiência variam em torno de seu valor médio, é preciso realizar a análise da variância.

Os pesquisadores envolvidos em tarefas desta natureza, via de regra, têm empregado o teste de Wilcoxon do *rank* de sinais, com índices emparelhados em uma única amostra (Ruggiero, 1998) e o teste Mann-Whitney da soma dos postos (Friedman & Stern, 1997). Em ambos os enfoques, os pesquisadores trabalham com amostras independentes. A presente pesquisa, porém, requer uma abordagem não-paramétrica que permita testar diferenças numa mesma amostra de UTD's, que tenham sido mensuradas mais de duas vezes. Isto caracteriza o caso de amostras relacionadas.

Na análise paramétrica da variância de observações contínuas, a influência das diferenças entre os blocos (UTD's, no caso) é removida pela mensuração desta fonte de variabilidade, antes de serem procedidas as comparações requeridas pela análise. Na pesquisa, esta "filtragem" é muito importante, pois não interessa a magnitude da variância dos índices entre as UTD's, mas sim entre os modelos de análise. Um teste não-paramétrico com esta característica é o de Friedman, em que a soma dos postos em todos os blocos é a mesma. Segundo Sprent (1990), a maior vantagem do ordenamento dentro dos blocos é que ele proporciona um método razoavelmente robusto para muitas formas de heterogeneidade da variância.

Para Sprent (1990), o teste de Friedman é uma extensão natural do teste de sinais de Wilcoxon e funciona muito bem para amostras grandes, que sejam relacionadas e cujos dados sejam ordenados. Para aplicar este teste na pesquisa, as linhas da matriz de índices

são ocupadas pelos diferentes modelos, e as colunas, pelas UTD's. Deste modo, o ordenamento é feito por coluna, separadamente, para cada UTD. Este procedimento produz Rk_{ji} , que é o posto atribuído ao índice de eficiência técnica estimado pelo j -ésimo modelo para a i -ésima UTD.

A forma da estatística de Friedman comumente usada é

$$QQF = \frac{n(l-1)(SQRK - FC)}{(STQRK - FC)} \quad (3.153)$$

onde: $SQRK = \left(\sum_{j=1}^l SRK_j^2 \right) / n;$

$$SRK_j = \sum_{i=1}^n RK_i \quad \text{é a soma dos postos das } n \text{ UTD's, segundo o } j\text{-ésimo modelo;}$$

$$FC = \frac{1}{4} n l (l+1)^2 \quad \text{é um fator de correção e}$$

$$STQRK = \sum_{j=1, i=1}^{l, n} RK_{ji}^2 \quad \text{é a soma dos quadrados de todos os postos da matriz de dados ordenados.}$$

Para n e l não muito pequenos, QQF tem aproximadamente distribuição de qui-quadrado com $l-1$ graus de liberdade.

Segundo Iman & Davenport *apud* SPRENT (1990), uma melhor aproximação para a estatística de Friedman é

$$FF = \frac{(n-1)(SQRK - FC)}{(STQRK - SQRK)} \quad (3.154)$$

que tem aproximadamente distribuição F com $(l-1)$ e $(n-1)(l-1)$ graus de liberdade. A referida estatística é utilizada para testar a seguinte hipótese nula:

$$H_0: \text{Não há diferenças entre os tratamentos (modelos)} \quad (3.155)$$

contra

$$H_1: \text{Pelo menos um tratamento tem postos diferentes.}$$

O teste de Friedman, conforme apresentado acima, indica significância global. Quando a hipótese (3.155) é rejeitada, deve-se executar o teste de Friedman para

comparações múltiplas. Através dele, a significância de comparações específicas é testada pelo exame das discrepâncias entre os postos dos diferentes tratamentos (dois a dois) ou entre combinações entre eles. Em comparações bivariadas, a magnitude das diferenças é considerada significativa quando ela excede a diferença significativa mínima, dada por

$$DSM = \left(t_{(n-1)(l-1)}(n.s) \right) \sqrt{\frac{2n(STQRK - SQRK)}{(n-1)(l-1)}} \quad (3.156)$$

onde $\left(t_{(n-1)(l-1)}(n.s) \right)$ é o valor crítico da distribuição t com $(n-1)(l-1)$ graus de liberdade, a um determinado nível de significância $(n.s)$. Desta forma, toda vez que

$$DSM > (SRK_j - SRK_{j'}), \quad j \neq j' \quad (3.157)$$

aceita-se a hipótese nula de nenhuma diferença estatisticamente significativa entre os índices de eficiência técnica estimados pelos modelos j e j' .

Mais informações sobre as distribuições dos índices de eficiência podem ser extraídas do exame de seus percentis e quartis. Também, é útil a visualização das tabelas de frequência relativa dos índices de eficiência e os seus respectivos histogramas.

3.5 PROGRAMAS COMPUTACIONAIS

A abordagem de funções de fronteira estocástica conta, até o presente, com dois *softwares* que tratam a sua estimação de maneira particular. Um deles, desenvolvido por Coelli (1996), é denominado FRONTIER (*a computer program for stochastic frontier production and cost function estimation*). É razoavelmente completo, porém, excessivamente automático na execução dos comandos que são bem específicos.

O outro, chamado LIMDEP (*limited dependent variable models*), construído por Greene (1995), é um completo pacote destinado a estimar e analisar modelos econométricos básicos e avançados. Além de um comando específico para funções de fronteira, ele oferece numerosas ferramentas de programação, incluindo uma versátil rotina de álgebra matricial e rotinas de otimização. Estes recursos possibilitam definir funções próprias de verossimilhança e adicionar novas especificações na lista dos modelos

contemplados pelo programa. Todos os resultados obtidos podem ser guardados e empregados em etapas posteriores.

São, também, de muita utilidade, outras ferramentas de programação como o editor, os comandos *looping* e de transformação de dados, e as facilidades para criar novos procedimentos. Com estes recursos pode-se construir aplicações próprias para novos modelos ou para a execução de uma grande variedade de cálculos, inclusive os envolvidos em testes estatísticos mais complicados, a construção de matrizes de covariância, etc...

O LIMDEP contém, também, um poderoso gerador de números e de amostras aleatórias para treze distribuições discretas e contínuas, incluindo a normal, truncada normal, Poisson, uniforme, binomial, logística e Weibull, dentre outras.

Com todos estes atrativos, preferir o LIMDEP frente ao FRONTIER, é uma decisão tranquila. Além disto, é irrecusável a orientação direta oferecida pelo criador do LIMDEP, quando da permanência do autor desta pesquisa em Nova York, em meados de 1997.

Por essas razões, todo o trabalho de geração da base de dados da pesquisa e de estimação e verificação dos modelos de análise é programado para processamento no LIMDEP. A análise comparativa dos resultados, apesar da sofisticação do teste de Friedman, envolve procedimentos simples, cujo desenvolvimento tem melhor fluidez em planilhas do Microsoft Excel²⁰.

RESULTADOS E INTERPRETAÇÃO

²⁰ Todos os programas e planilhas elaborados podem ser fornecidos mediante solicitação.

Na realização da parte aplicada da pesquisa, foram utilizados os programas computacionais elaborados de acordo com a metodologia definida no capítulo 3.

Os relatórios de saída dos referidos programas são deveras extensos, pois envolvem matrizes de dados com 161 linhas e uma gama bastante variada de procedimentos para os 6 modelos analisados. Por isso, inclui-se neste capítulo apenas as tabelas com os resultados principais. A listagem completa dos resultados está disponível para fornecimento mediante solicitação.

4.1 CONSTRUÇÃO DOS DADOS AMOSTRAIS

Para construção dos dados amostrais da pesquisa, foi operacionalizada a metodologia de geração e seleção de amostras aleatórias apresentada na seção 3.3.

Dentre as amostras geradas, foram descartadas as que apresentaram modelos com valores positivos de m_3 nas funções de produção (e negativos nas de custo e de produção inversa) e valores negativos para s_V^2 ou s_U^2 . Com este procedimento evitou-se a ocorrência de erros do Tipo I e do Tipo II (ver detalhes na seção 3.1.3.3).

O processo de geração de amostras aleatórias foi encerrado no momento em que a soma de $\sqrt{b_1}$, referente aos seis modelos de análise, em cada nova amostra gerada, cessou de aumentar com relação aos totais anteriores. Quando isto aconteceu, já haviam sido produzidas mais de 1.500 amostras aleatórias. Desta forma, foi escolhida como amostra final (única), a amostra que atingiu a maior soma total de $\sqrt{b_1}$.

4.2 ESTIMAÇÃO E VERIFICAÇÃO DOS MODELOS DE ANÁLISE

Na estimação dos modelos de análise e dos respectivos índices de eficiência técnica para cada UTD, foram empregados os dados da amostra final. As estimativas dos coeficientes de regressão, dos valores de $\sqrt{b_1}$, s^2 , s_U^2 , s_V^2 , g , g^* e R^2 e dos escores de eficiência estão contidas na tabela 4.1.

Tabela 4.1 - Estimativas dos parâmetros das funções de fronteira estocástica,

segundo cada modelo.

Regressor	Parâmetro	PAI	PAC	CP	PI	DP	PM
	$b_{0\text{col}/\text{ccca}}$	1,4009 (0,1426) (0,1575)	1,6757 (0,1766) (0,1921)	1,2961 (0,2013) (0,1787)	-3,1062 (0,4873) (0,4246)	-0,9050 (0,1917)	-2,2500 (0,1344)
$\ln x_{1i}$	b_1	0,4053 (0,0220) (0,0237)	0,4183 (0,0272) (0,0281)			-0,3681 (0,0295)	1,1761 (0,0463)
$\ln x_{2i}$	b_2	0,4184 (0,0133) (0,0161)	0,4355 (0,0162) (0,0196)			-0,4064 (0,0179)	1,2945 (0,0097)
$\ln y_{1i}^{\text{pai}}$	a_1^{pai}	0,0934 (0,0203) (0,0237)					
$\ln y_{1i}$	a_1			0,2606 (0,0258) (0,0271)	0,5010 (0,0693) (0,0776)		0,8408 (0,0068)
$\ln y_{2i}$	a_2			0,6093 (0,0349) (0,0299)	1,3188 (0,0897) (0,0721)		1,9357 (0,0095)
$\ln(w_{1i}/w_{2i})$	h_1/k			0,5322 (0,0121) (0,0108)			
$\ln(y_{1i}/y_{2i})$	a_1^{dp}					0,3264 (0,0173)	
$\ln x_{1i}^{\text{pi}}$	b_1^{pi}				-1,0417 (0,0622) (0,0575)		
	$\sqrt{b_1}$	-0,3080	-0,8988	0,7887	0,8259	-0,7381	-0,3622
	s^2	0,0280	0,0647	0,0794	0,6284	0,0689	0,6946
	s_u^2	0,0196	0,0631	0,0748	0,5995	0,0637	0,5147
	s_v^2	0,0084	0,0016	0,0046	0,0288	0,0051	0,1799
	g	0,6989*	0,9751	0,9425	0,9541	0,9258	0,7410*
	g^\bullet	0,4575	0,9343	0,8563	0,8831	0,8193	0,5098
	R^2	0,8642	0,8201	0,9664	0,8423	0,8745	0,8161

* estatisticamente significativo ao nível de 5%.

Nota: todos os demais coeficientes são significantes ao nível de 1%.

Na tabela 4.1, à exceção do modelo PM, os primeiros valores entre parênteses são as estimativas de OLS dos erros-padrão dos coeficientes de regressão associados às variáveis explicativas. Os erros-padrão dos interceptos, excetuado o do modelo PM, foram estimados por COLS. No modelo de fronteira de produção múltipla, os erros-padrão de todos os coeficientes de correlação canônica foram estimados por CCA pelo

método *jackknife*. A estimativa do erro-padrão do coeficiente de interseção do modelo PM foi calculada por CCCA.

Os segundos valores entre parênteses referem-se às estimativas de MLE dos erros-padrão dos coeficientes. Somente para os modelos de fronteira de distância-produto e de fronteira de produção múltipla, não se obteve as referidas estimativas. No primeiro caso, porque não foi possível abrir o LIMDEP para introduzir as transformações a que a variável dependente foi submetida. No segundo caso, porque o comando utilizado opera exclusivamente com uma única variável dependente.

As inexpressivas diferenças entre os valores dos erros-padrão dos coeficientes de regressão confirmam as conclusões de Olson et al. (1980) sobre a correspondência da estimação por COLS e por MLE.

Quanto à significância estatística, constata-se que apenas dois dos coeficientes estimados não se mostram significantes a 1%. Isto ocorre no teste da existência de uma fronteira para os modelos PAI e PM, em que a hipótese $H_0: \mathbf{g} = 0$ é rejeitada só ao nível de 5%. Estes testes são realizados por meio da estatística $\sqrt{b_1}$, cujos valores críticos tabelados, para uma amostra de tamanho 161, são 0,298 e 0,430 para 5 e 1%, respectivamente. Por isso, pode-se considerar todas as fronteiras estimadas como sendo estatisticamente muito boas.

A significância dos coeficientes de determinação não foi testada porque eles provêm das estimativas iniciais por OLS. Porém, é importante reconhecer que o elevado poder de explicação daquelas equações (R^2 nunca inferior a 0,81) é estatisticamente salutar para as respectivas funções de fronteira.

A hipótese de seminormalidade dos U_i 's foi avaliada, preliminarmente, através da observação do formato da distribuição de U_i em cada modelo. Os referidos formatos sugerem distribuições bem aproximadas da curva normal truncada, como se pode ver nos histogramas inseridos nos apêndices A a F. Após a referida inspeção visual, foram realizados testes de aderência dos U_i 's às distribuições normal truncada e exponencial (tabela 4.2).

Tabela 4.2 - Teste da aderência de U_i pelo qui-quadrado.

Modelo	H_0	C_{calc}^2	Decisão
PAI	U_i 's são normais truncados	133,858	Rejeita-se H_0 a 1%
	U_i 's são exponenciais	152,487	Rejeita-se H_0 a 1%
PAC	U_i 's são normais truncados	111,276	Rejeita-se H_0 a 1%
	U_i 's são exponenciais	132,213	Rejeita-se H_0 a 1%
CP	U_i 's são normais truncados	128,640	Rejeita-se H_0 a 1%
	U_i 's são exponenciais	70,1812	Rejeita-se H_0 a 1%
PI	U_i 's são normais truncados	31,0798	Rejeita-se H_0 a 1%
	U_i 's são exponenciais	57,0548	Rejeita-se H_0 a 1%
DP	U_i 's são normais truncados	94,1848	Rejeita-se H_0 a 1%
	U_i 's são exponenciais	70,6185	Rejeita-se H_0 a 1%
PM	U_i 's são normais truncados	41,3052	Rejeita-se H_0 a 1%
	U_i 's são exponenciais	49,0034	Rejeita-se H_0 a 1%

Os resultados sintetizados na tabela 4.2 evidenciam altos níveis de significância na rejeição da hipótese nula. Por conseguinte, pode-se aceitar, com boa margem de confiabilidade, que os U_i 's de todos os modelos analisados seguem, efetivamente, distribuições seminormais.

No que concerne às verificações específicas, a pesquisa ficou restrita à hipótese de minimização de custos, à averiguação da existência de um insumo dominante e à avaliação da qualidade estatística do ajuste canônico da função de produção múltipla e de suas condições de primeira e de segunda ordens.

Nos dois primeiros casos, é suficiente recorrer às expressões contidas em (3.148) para deduzir facilmente que, por construção, o dispêndio com qualquer um dos dois insumos representa, para cada uma das UTD's amostradas, exatamente a metade do seu custo total. Portanto, a suposição de exata minimização de custos (3.112) é aceita, porque a referida relação é sempre igual à unidade para qualquer UTD. Nestas condições, também é óbvia a inexistência de um insumo dominante. Por esta razão, estimou-se, no caso da fronteira de produção inversa, a variante com os insumos agregados em um índice.

Por seu turno, a significância global do ajustamento canônico foi avaliada através das estatísticas de qui-quadrado (Bartlett) e F (Rao). A síntese de todas as hipóteses testadas e as respectivas decisões estão na tabela 4.3.

Tabela 4.3 - Teste da significância global do ajustamento canônico do modelo PM.

H_0	C_{calc}^2	F_{calc}	Decisão
$\sum_{yx} = 0 \quad (r_1 = r_2 = 0)$	237,278	$C_4^2(0,01)=13,28$	Rejeita-se H_0
$\sum_{yx} = 0 \quad (r_1 = r_2 = 0)$		88,2278	$F_{4,316}(0,01)=3,32$, Rejeita-se H_0
$r_1 \neq 0 \quad (r_2 = 0)$	0,1914	$C_1^2(0,01)=6,63$	Aceita-se H_0
$r_1 \neq 0 \quad (r_2 = 0)$		0,0477	$F_{4,316}(0,01)=3,32$, Aceita-se H_0

Em todas as hipóteses testadas na tabela 4.3, evidencia-se a grande significância estatística individual do coeficiente de correlação máximo para o primeiro par de variáveis canônicas, ou seja, de r_1 .

Avaliando os resultados do modelo PM, frente a alguns requisitos da teoria econômica, verificou-se que eles ocorrem conforme esperados. Todas as elasticidades e produtividades marginais são positivas. No primeiro caso, as referidas estimativas estão na tabela 4.4 e, no segundo, podem ser fornecidas mediante pedido.

Tabela 4.4 - Estimativas das elasticidades marginais de cada produto com relação a cada insumo, segundo o modelo PM

Elasticidade	Estimativa
$EMg(y_1, x_1)$	1,3987
$EMg(y_1, x_2)$	1,5396
$EMg(y_2, x_1)$	0,6076
$EMg(y_2, x_2)$	0,6687

Também estão disponíveis para consulta, os valores das estimativas das taxas marginais de transformação entre os dois produtos e de substituição entre os dois insumos. A confirmação geométrica destas características, algebricamente determinadas, está no

formato decrescente das relações entre os dois produtos e entre os dois insumos, o qual pode ser visualizado nos gráficos constantes dos apêndices G e H, respectivamente. Dos referidos gráficos, infere-se, também, que as curvas de transformação e de substituição são convexas para a origem.

Estas constatações geométricas sobre a convexidade são confirmadas, algebricamente, pelos valores de $\text{Conv}(x_{1i}, x_{2i})$ e de $\text{Conv}(y_{1i}, y_{2i})$. Na listagem completa, que pode ser obtida mediante requisição, os referidos valores são positivos para todas as UTD's. Isto reflete a presença de economias de escala que são bem visíveis nas elasticidades marginais. Desta forma, as curvas de iso-receita são convexas à origem, obedecendo o que está previsto na teoria para o caso de preços variáveis.

Em conclusão, com base no conjunto de resultados algébricos e geométricos obtidos, pode-se afirmar que a função estimada de fronteira estocástica de produção múltipla atende satisfatoriamente as condições de primeira e de segunda ordens.

4.3 ANÁLISE COMPARATIVA

A análise comparativa dos índices de eficiência técnica, referentes a cada modelo estimado e testado na seção anterior, foi consumada mediante o uso de uma planilha do Microsoft Excel, elaborada de conformidade com a seção 3.4.

Inicialmente, calculou-se o coeficiente de correlação ordinal de Spearman entre os escores das UTD's segundo cada modelo, dois a dois, mediante o emprego de (3.151). Para testar a hipótese de que as referidas correlações fossem nulas, comparou-se o valor de $t_{j,j'}(\text{calc})$ (obtido por 3.152) com os respectivos valores críticos tabelados. Estes resultados estão expostos na tabela 4.5, onde os valores entre parênteses são os correspondentes valores calculados de t .

Tabela 4.5 - Matriz de correlação ordinal (de Spearman) entre os índices de eficiência técnica estimados em cada modelo.

	PAI	PAC	CP	PI	DP	PM
PAI	1,0000					
PAC	0,8464 (20,0424)	1,0000				
CP	0,7203 (13,0971)	0,6811 (11,7325)	1,0000			
PI	0,4735 (6,7803)	0,3871 (5,2943)	0,5840 (9,0737)	1,0000		
DP	0,8750 (22,7935)	0,7606 (14,7755)	0,7686 (15,1497)	0,4818 (6,9346)	1,0000	
PM	0,4972 (7,2274)	0,2260 (2,9263)	0,7294 (13,4474)	0,4586 (6,5090)	0,5454 (8,2064)	1,0000

Observa-se na tabela 4.5 que todos os coeficientes de correlação são estatisticamente significantes a 1% (a este nível o valor crítico de t , para 159 graus de liberdade, é 2,326).

Na sequência da análise comparativa, testou-se a hipótese nula global da não existência de diferença, estatisticamente significativa, entre os índices produzidos pelos diferentes modelos estimados. Para isto, tomou-se por base o teste não-paramétrico de Friedman, cuja estatística foi computada por (3.154). O valor desta estatística (167,3983) comparado com o valor crítico da distribuição F com 5 graus de liberdade no denominador e 800 graus de liberdade no numerador, ao nível de significância de 1% (9,02), resultou na rejeição de H_0 . Isto implica em que pelo menos um dos modelos apresenta diferenças estatisticamente significantes com relação aos índices produzidos, devendo-se, por isso, examinar a significância entre comparações específicas.

Essas verificações foram realizadas por meio do teste de Friedman para comparações múltiplas, em cuja construção empregou-se a estatística da diferença da soma dos postos ($SRK_j - SRK_{j'}, j \neq j'$), utilizando os mesmos valores de SRK_j calculados no teste anterior. Os valores da referida estatística, para cada sub-conjunto de dois modelos, estão na tabela 4.6.

Tabela 4.6 - Matriz das somas dos postos (entre parênteses) e das diferenças das somas para cada modelo.

	PAI	PAC	CP	PI	DP	PM
PAI	(902)					
PAC	301	(601)				
CP	313*	12	(589)			
PI	606*	305*	293	(296)		
DP	274	27	39	332*	(628)	
PM	538*	237	225	68	264	(364)

Na tabela 4.6, nos casos assinalados com um asterisco, a hipótese $H_0: ET_j = ET_{j'}, j \neq j'$, é rejeitada porque as respectivas diferenças são maiores do que a diferença significativa mínima (*DSM*). Esta diferença, calculada por (3.156), é igual a 302,6238 para um valor crítico de t com 800 graus de liberdade, ao nível de significância de 1%. Os modelos PAI e PI apresentam três situações de rejeição da hipótese nula. Nos demais modelos, a não aceitação de H_0 se dá em apenas uma comparação. Nos modelos PAC e DP, ela acontece no confronto com o modelo PI, e nos modelos CP e PM, no cotejamento com o modelo PAI.

Este conjunto de decisões é suficiente para se inferir que os índices estimados pelos modelos PAI e PI diferem substancialmente dos demais. Em outras palavras, pode-se dizer que eles provêm de populações diferentes das que deram origem aos índices dos modelos PAC, CP, DP e PM. As diferenças entre os índices destes quatro modelos, por sua vez, não demonstram significância estatística relevante. Eles são originários da mesma população.

4.4 INTERPRETAÇÃO DOS RESULTADOS

Todos os resultados, obtidos na fase de estimação e verificação dos modelos de análise, são altamente significativos, sob os pontos de vista estatístico e econométrico. Isto não é surpreendente, porque foi empregada uma tecnologia de produção bem definida (CET-CD) na geração das variáveis aleatórias e selecionou-se, dentre as amostras

geradas, a amostra com as fronteiras estocásticas de maior significância global. Além disso, a ocorrência de altos níveis de significância também é devida às especificações adotadas para garantir comparabilidade.

Somente os modelos CP e PM foram examinados quanto ao atendimento de requisitos da teoria econômica. Esta providência não foi necessária nos demais modelos. O modelo de produção inversa segue as especificações da fronteira de custo, e os outros constituem versões da bem conhecida função de produção Cobb-Douglas.

As condições requeridas para a função de custo ser efetivamente dual da função de produção foram todas contempladas na construção dos dados amostrais e no conjunto de especificações. Deste modo, atendeu-se as exigências de minimização de custos, de homogeneidade linear, de monotonicidade e de convexidade.

O modelo de produção múltipla, por ser não-trivial, demandou uma investigação completa a respeito do cumprimento das condições de primeira e de segunda ordens. Para este modelo, computou-se as estimativas das elasticidades e produtividades marginais, das taxas marginais de transformação e de substituição e das condições de convexidade, verificando-se que as referidas grandezas apresentaram os sinais e as magnitudes esperados.

Uma vez atendidos todos os requisitos estatísticos, econométricos e de teoria econômica necessários, procedeu-se à comparação dos índices de eficiência técnica, por intermédio de análises não-paramétricas de correlação e de variabilidade.

No primeiro tipo de análise, verificou-se que todos os coeficientes de correlação ordinal de Spearman são altamente significantes, ainda que a correlação entre os modelos PAC e PM tenha acusado um valor demasiado diminuto. A baixa correlação entre os índices desses dois modelos deve-se às grandes discrepâncias observadas nos índices de umas poucas UTD's. Estas discrepâncias são explicadas, em boa parte, pelas expressivas diferenças ocorridas nas quantidades empregadas dos dois insumos e pela excessiva diferenciação dos coeficientes de regressão das referidas variáveis, de um modelo para outro.

Na análise de variância através dos testes de Friedman com comparações múltiplas, concluiu-se que os modelos PAC, CP, DP e PM, em média, classificam as UTD's em postos que não diferem substancialmente entre si, ao nível de significância de 1%. Nestes testes, o ordenamento dos escores de eficiência é feito por UTD, segundo os modelos, enquanto que na análise de correlação, ele se dá por modelo, segundo as UTD's. Por isso, pode ocorrer a aceitação da hipótese nula pelo teste de Friedman, na presença de um coeficiente de correlação bastante baixo, como aconteceu com os índices PAC e PM. Ou, a rejeição de H_0 , a despeito de um coeficiente de correlação expressivo, como no caso da comparação entre os índices PAI e CP.

Na tabela 4.5, observa-se que os índices do modelo CP são os que melhor se correlacionam com todos os outros, caracterizando, assim, um perfil de referência (*benchmark*). Ordenando os índices dos demais modelos significantes, segundo este perfil, pode-se avaliar, comparativamente, o desempenho de cada um. Estes resultados, para as 15 UTD's mais eficientes e para as 15 mais ineficientes, estão expostos na tabela 4.7.

Tabela 4.7 - Discriminação das 15 UTD's mais eficientes e das 15 mais ineficientes, segundo cada modelo.

UTD's	CP	DP	PAC	PM
MAIS EFICIENTES				
81	x			x
146	x	x	x	
59	x			x
62	x			x
60	x	x	x	x
57	x			x
69	x	x	x	
8	x		x	
32	x			x
157	x	x	x	
86	x			x
73	x			x
75	x	x		x
82	x			
92	x	x	x	
TOTAL	15	6	6	9
				(continua)

Tabela 4.7 - Discriminação das 15 UTD's mais eficientes e das 15 mais ineficientes, segundo cada modelo.
(continuação)

UTD's	CP	DP	PAC	PM
MAIS INEFICIENTES				
78	x	x		x
70	x			x
130	x	x	x	
91	x			x
67	x			x
158	x		x	
1	x			x
71	x	x		x
61	x	x		x
2	x	x		x
85	x	x		x
126	x	x	x	
127	x	x	x	
95	x	x	x	
141	x	x	x	
TOTAL	15	10	6	9

Nota: os números que constam na primeira coluna referem-se aos números de ordem da UTD na amostra de dados.

Na tabela 4.7, constata-se que o modelo novo (PM) identifica 9 das 15 UTD's sabidamente eficientes, enquanto que os modelos DP e PAC conseguem distinguir somente 6 das referidas unidades. Na classificação das 15 unidades mais ineficientes, o modelo PM suplanta de novo o modelo PAC e se equipara ao modelo DP.

Outra constatação interessante é que o modelo PM não concentrou os índices nos percentis superiores como aconteceu com os modelos CP e DP, conforme se pode ver na figura 4.1.

Em recente pesquisa desenvolvida por Ruggiero (1998), este aspecto foi detectado, porém no confronto do método CCA com o método DEA, em um contexto determinístico. Ruggiero interpretou este fenômeno como uma tendência do método CCA a superestimar a ineficiência.

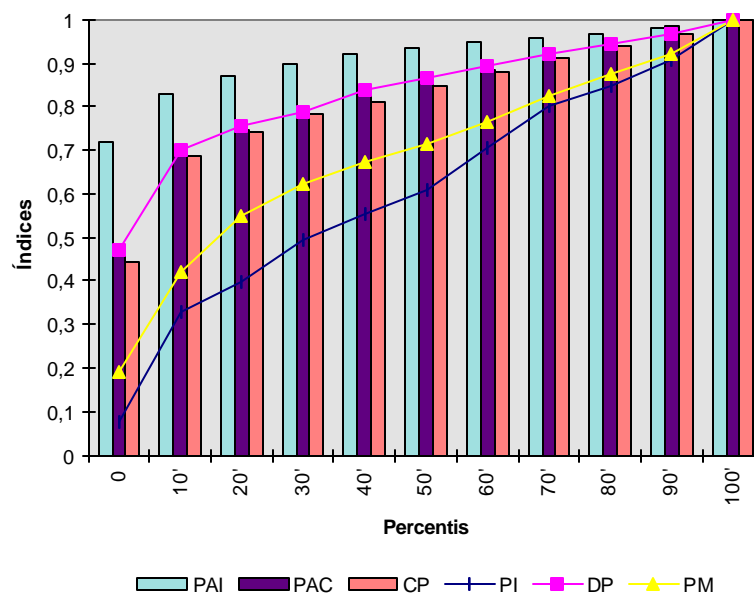


Figura 4.1 - Percentis das distribuições dos índices de eficiência estimados pelos modelos de análise.

No caso da presente pesquisa, em um ambiente estocástico, revive-se o mesmo fenômeno, porém com uma característica a mais. Comparativamente aos outros modelos, o modelo PM distribuiu melhor os índices entre todos os percentis, principalmente nos iniciais. O resultado disto é que se obtém informações mais detalhadas sobre as UTD's mais ineficientes. O modelo PI também produziu escores com esta característica, porém com maior intensidade. Enquanto ele acusou 20% das UTD's com índices de valor até 0,3985, o modelo PM registrou um índice igual a 0,5480, no vigésimo percentil.

O método econométrico vem sendo aplicado cada vez mais na análise de eficiência produtiva, principalmente desde o surgimento do modelo de fronteira estocástica de produção, em 1977.

Uma crítica comum à análise econométrica de eficiência atacava a sua inabilidade na avaliação de tecnologias com múltiplos produtos. Naquela época, a abordagem limitava-se a fronteiras de produção agregada, em que os vários produtos são juntados em uma só variável dependente. A partir de fins da década de 80, começou-se a utilizar fronteiras de valores duais (custo, receita e lucro). Um pouco mais tarde, as opções incluíam a fronteira de produção inversa e a fronteira de distância-produto. Todos estes modelos são uniequacionais, uma vez que no caso multiequacional eles são demasiado precoces e apresentam sérias dificuldades de estimação.

Nesse processo evolutivo, os diferentes modelos têm sido aplicados isoladamente e sem muita preocupação com as implicações decorrentes de especificações inadequadas e da falta de atendimento a requisitos estatísticos, econométricos e de teoria econômica. Sob este prisma, a presente pesquisa é uma primeira contribuição, ainda que tal questão não figure como seu objetivo principal. Ou seja, a análise comparativa levada a efeito destina-se, exclusivamente, a subsidiar a proposição e a verificação do novo modelo frente aos demais. Não obstante, alguns resultados inerentes aos modelos vigentes merecem ser destacados.

Os escores de eficiência estimados pelos modelos de produção agregada mostram-se excessivamente diferenciados dos demais, principalmente no caso da agregação por índice (modelo PAI). Neste modelo, encontra-se correspondência estatisticamente aceitável apenas com os índices produzidos através da fronteira de distância-produto (modelo DP). Já o modelo de fronteira de produção agregada por correlação canônica (PAC), demonstra boa similaridade com o modelo de fronteira de custo (CP), com o modelo DP e com o modelo novo de fronteira de produção múltipla (PM). Assim, na impossibilidade de utilizar um destes três modelos, pode-se lançar mão do modelo PAC.

Por sua vez, o uso do modelo de fronteira de produção inversa (PI) está condicionado à existência de um insumo dominante, podendo-se, também, utilizar um índice agregado dos insumos. Este último procedimento, contudo, pode produzir resultados inadequados, conforme se verifica na pesquisa, onde tais índices assemelham-se apenas aos provindos dos modelos CP, DP e PM (novo), apresentando correlações bivariadas que não passam de 0,5840. Além disso, há que se ponderar que é bem difícil justificar a preferência pelo recurso da agregação dos insumos, ao invés do emprego da fronteira de custo.

Já o modelo de fronteira de distância-produto revela-se como um caminho deveras atraente, desde que a função seja não-decrescente, positiva, linearmente homogênea e convexa em Y , e decrescente em X . Com estas condições garantidas na pesquisa, os índices estimados por este modelo são considerados correspondentes aos obtidos através dos modelos PAC, CP e PM. A função dual da função distância-produto é a função receita que, por este motivo, não é investigada. Ademais, não foi encontrada na literatura nenhuma referência sobre o desenvolvimento teórico-aplicado desta função na análise de eficiência técnica. A fronteira de lucros também não é examinada por razões parecidas.

Em consequência destas últimas constatações, o contexto metodológico uniequacional do enfoque dual da fronteira de produção, na prática, fica confinado ao âmbito da fronteira de custo. Na pesquisa, os índices estimados pelo modelo CP estão fortemente correlacionados com todos os outros, inclusive com os gerados pelos modelos PAI e PI. O único revés estatístico experimentado por aqueles índices é a rejeição de H_0 no teste de Friedman, quando comparados com os índices do modelo PAI. É importante frisar que a boa performance do modelo CP é alcançada a par do atendimento dos requisitos de minimização de custos, homogeneidade linear, monotonicidade e convexidade nos preços dos insumos.

Quanto ao desempenho estatístico, o modelo novo de fronteira de produção múltipla (PM) também sofre apenas um fracasso no teste de Friedman e apresenta correlações estatisticamente importantes, porém com magnitudes invariavelmente inferiores às registradas pelo modelo CP. No caso da comparação com o modelo PAC, ela é

demasiado pequena, em decorrência de grandes discrepâncias em alguns postos. Não obstante, não há como concluir que os índices dos dois modelos não sejam assemelhados.

Portanto, com base nas análises não-paramétricas de correlação ordinal e de variância levadas a cabo na pesquisa, pode-se concluir que os índices estimados pelos modelos PAC, CP, DP e PM são provenientes da mesma população e tem correlações estatisticamente aceitáveis entre si (dois a dois). Na tomada destas decisões, todos os testes de hipóteses envolvidos são realizados a níveis de significância de 1%. Especificamente, o modelo ora proposto é mais sensível do que os modelos PAC e DP na definição das UTD's mais eficientes, tomando o modelo CP como *benchmark*. Na determinação das unidades mais ineficientes, ele demonstra a mesma sensibilidade do modelo DP e continua superando o modelo PAC.

Os resultados alcançados permitem concluir, também, que a utilização do modelo PM é sempre viável e, sob certas circunstâncias, pode tornar-se a única alternativa possível para a análise econométrica uniequacional de eficiência técnica. Isto pode acontecer em cenários que conjuguem as seguintes situações:

- a) quando existe produção conjunta dos múltiplos produtos em um único processo produtivo, o emprego da fronteira de produção agregada é inadequado;
- b) na falta de um insumo dominante, a estimação da fronteira de produção inversa, via agrupamento dos insumos, é problemática;
- c) a dificuldade maior, no caso da fronteira de distância-produto, pode residir no não cumprimento do requisito de homogeneidade;
- d) quando a especificação da fronteira de custo não atende todos os requisitos exigidos, ela pode não ser seguramente dual ou refletir a eficiência técnica e alocativa, conjuntamente. No primeiro caso, ela é inadequada. No segundo, a sua decomposição implica na convivência com problemas de estimação ainda não suficientemente resolvidos e, ainda,
- e) a falta de dados sobre preços pode ser outro empecilho ao uso das fronteiras de custo.

Assim sendo, a fronteira estocástica de produção múltipla nasce com importância substancialmente valorizada: o modelo PM é bem mais flexível do que os demais e é uma alternativa sempre viável (que pode tornar-se única). Além disso, o seu desenvolvimento

ensejou a consecução de realizações pioneiras. É o caso dos programas computacionais elaborados de forma integrada, da construção da amostra aleatória única, da estimação de todas as fronteiras sob especificações adequadas e do registro, na literatura econométrica brasileira, da estimação de uma função de distância-produto.

Apesar de validado para especificações bem definidas, o novo modelo é passível de algumas extensões. A sua expansão para mais de dois produtos e para mais de dois insumos, não é nem um pouco complicada. Por exemplo, variáveis sobre a qualidade dos produtos podem ser incluídas no vetor de produção; sobre a qualidade dos fatores, no vetor de insumos; sobre a utilização da capacidade instalada, no vetor de insumos; e assim por diante.

Não obstante os aspectos inovadores e os benefícios frente à tecnologia vigente, a pesquisa ressen-te-se da inexistência de fórmulas para cálculo dos erros-padrão dos parâmetros da função de produção múltipla. A superação desta limitação abre oportunidade para importantes pesquisas futuras. No caso da estimação por CCCA, pode-se trabalhar no desenvolvimento teórico dos erros-padrão de todos os coeficientes canônicos. No caso da estimação por MLE, pode-se almejar a construção de um algoritmo de otimização de funções (de produção de fronteira estocástica) de verossimilhança com mais de uma variável dependente. A raiz quadrada dos elementos da diagonal da matriz de informação, oriunda da iteração final do referido algoritmo, fornece os erros-padrão aproximados das estimativas de MLE.

Os resultados obtidos na presente pesquisa provêm da estimação de modelos com especificações bem delimitadas quanto à definição das variáveis, ao número de equações, ao método de estimação, à forma funcional, à estocasticidade da fronteira, à distribuição dos resíduos, à natureza e tamanho da amostra de dados e ao atendimento de requisitos estatísticos, econométricos e da teoria econômica. A execução de outros trabalhos, sob diferentes especificações, abre imensas possibilidades para pesquisas futuras. Essas novas investigações, ao incorporar o conhecimento produzido na atual, podem ser melhor avaliadas por meio de simulações do tipo Monte Carlo.

REFERÊNCIAS BIBLIOGRÁFICAS

- AIGNER, D.J., CHU, S.F. *On estimating the industry production function*. **American Economic Review**, v.58, p.826-839, 1968.
- AIGNER, D., LOVELL, C.A.K., SCHMIDT, P. *Formulation and estimation of stochastic frontier production function models*. **Journal of Econometrics**, v.6, p.21-37, 1977.
- ARNOLD, L., BARDHAN, I.R., COOPER, W.W. at al. *New uses of DEA and statistical regressions for efficiency evaluation and estimation - with an illustrative application to public secondary schools in Texas*. In: **Annals of Operations Research**, 66, 1996, USA. Anais ..., p. 255-277.
- BATTESE, G.E., COELLI, T.J. *Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data*. **Journal of Econometrics**, v.38, p.387-399, 1988.
- BATTESE, G.E., CORRA, G.S. *Estimation of a production frontier model: with application to the pastoral zone of Eastern Australia*. **Australian Journal of Agricultural Economics**, v.21, n.3, p.169-179, 1977.
- BAUER, P.W. *Recent developments in the econometric estimation of frontiers*. **Journal of Econometrics**, v.46, p.39-56, 1990.
- BRAGA, H.C., ROSSI, J.W. *Mensuração da eficiência técnica na indústria brasileira: 1980*. **Revista Brasileira de Economia**, v.40, n.1, p.89-118, 1986.
- COBB, S., DOUGLAS, P. *A theory of production*. **American Economic Review**, v.18, p.139-165, 1928.
- COELLI, T. *Estimators and hypothesis tests for a stochastic frontier function: a Monte Carlo analysis*. **The Journal of Productivity Analysis**, v.6, p.247-268, 1995.

- _____. *A guide to FRONTIER version 4.1: a computer program for stochastic frontier production and cost function estimation*. Austrália: **University of New England, Centre for Efficiency and Productivity Analysis, Working Paper**, n.7, 1996.
- _____. [Standard errors of canonical coefficient] USA: University of New England, 22 set. 1999. Informação por Correio Eletrônico. Rede EPS. User name T.J.Coelli@reading.ac.uk.
- COSTA NETO, P.L.O. **Estatística**. São Paulo: Edgard Blücher, 1977. 264 p.
- CHIZMAR, J.F., ZAK, T.A. *Canonical estimation of joint educational production functions*. **Economics of Education Review**, v.3, n.1, p.37-43, 1984.
- CHRISTENSEN, L.R., GREENE, W.H. *Economies of scale in US electric power generation*. **Journal of Political Economy**, v.84, n.4, p.655-676, 1976.
- D'AGOSTINO, R.B., PEARSON, E.S. *Tests for departure from normality: empirical results for the distribution of b_2 and $\sqrt{b_1}$* . **Biometrika**, v.60, p.613-622, 1973.
- DEBREU, G. *The coefficient of resource utilization*. **Econometrica**, v.19, n.3, p.273-292, 1951.
- DIEWERT, W.E. *Functional forms for revenue and factor requirements functions*. **International Economic Review**, v.15, n.1, p.119-130, 1974.
- FÄRE, R., GROSSKOPF, S., LOVELL, C.A.K. **Production Frontiers**. USA: Cambridge University Press, 1994. 296 p.
- FÄRE, R., PRIMONT, D. **Multi-output production and duality: theory and applications**. USA: Kluwer Academic, 1995. 172 p.
- FARRELL, M.J. *The measurement of productive efficiency*. **Journal of the Royal Statistical Society**, v.120, parte III, p.253-281, 1957.
- FORSUND, F.R., LOVELL, C.A.K., SCHMIDT, P. *A survey of frontier production functions and of their relationship to efficiency measurement*. **Journal of Econometrics**, v.13, p.5-25, 1980.
- FRIEDMAN, L., SINUANY-STERN, Z. *Scaling units via the canonical correlation analysis in the DEA context*. **European Journal of Operational Research**, v.100, p.629-637, 1997.
- GATHON, H.-J., PERELMAN, S. *Measuring technical efficiency in european railways: a panel data approach*. **The Journal of Productivity Analysis**, v.3, p.135-151, 1992.

GATHON, H.J., PESTIEAU, P. *Decomposing efficiency into its managerial and its regulatory components: the case of european railways*. **European Journal of Operational Research**, v.80, p.500-507, 1995.

GYIMAH-BREMPONG, K. [Standard errors of canonical coefficient] USA: University of South Florida, 10 out. 1999. Informação por Correio Eletrônico. Rede EPS. User name KGyimah@coba.usf.edu.

GYIMAH-BREMPONG, K., GYAPONG, A.O. *Characteristics of education production functions: an application of canonical regression analysis*. **Economics of Education Review**, v.10, n.1, p.7-17, 1991.

GREENE, W.H. *On the estimation of a flexible frontier production model*. **Journal of Econometrics**, v.13, p.101-115, 1980.

_____. *Maximum likelihood estimation of stochastic frontier production models*. **Journal of Econometrics**, v.18, p.285-289, 1982.

_____. *A gamma distributed stochastic frontier model*. **Journal of Econometrics**, v.46, p.141-163, 1990.

_____. **Econometric Analysis**. EUA: Maxwell Macmillan, 1991. 783 p.

_____. *Frontier Production Functions*. USA: **New York University, Department of Economics, Working Paper**, EC-93-20, 104 p., 1993a.

_____. **The measurement of productive efficiency: techniques and applications**. USA: Oxford University Press, 1993b. 426 p. Cap. I2, p.68-119: *The econometric approach to efficiency analysis*.

_____. **LIMDEP version 7.0 user's manual**. USA: Econometric Software, 1995. 850 p.

GROSSKOPF, S. [Distance functions] USA: Southern Illinois University, 23 out. 1998. Informação por Correio Eletrônico. Rede EPS. User name Shawna.Grosskopf@orst.edu.

GROSSKOPF, S., HAYES, K. *Local public sector bureaucrats and their input choices*. **Journal of Urban Economics**, v.33, p.151-166, 1993.

HUGHES, M.D. *A stochastic frontier cost function for residential child care provision*. **Journal of Applied Econometrics**, v.3, p.203-214, 1988.

- JONDROW, J., LOVELL, C.A.K., MATEROV, I.S. et al. *On the estimation of technical inefficiency in the stochastic frontier production function model*. **Journal of Econometrics**, v.19, p.233-238, 1982.
- JOHNSON, R.A., WICHERN, D.W. **Applied multivariate statistical analysis**. 2 ed. USA: Prentice Hall, 1988.
- JUDGE, G.G., HILL, R.C., GRIFFITHS, W.E., LÜTKEPOHL, H., et al. **Introduction to the theory and practice of econometrics**. USA: John Wiley & Sons, 1988. 1.024 p.
- KARLSON, S.H. *Multiple-output production and pricing in electric utilities*. **Southern Economic Journal**, v.53, n.1, p.73-86, 1986.
- KOPP, R.J., DIEWERT, W.E. *The decomposition of frontier cost function deviations into measures of technical and allocative efficiency*. **Journal of Econometrics**, v.19, 319-331, 1982.
- KOOPMANS, T.C. **Activity analysis of production and allocation** USA: John Wiley & Sons, 1951, Cowles Commission for Research in Economics, Monograph n.13: *An analysis of production as an efficient combination of activities*
- KUMBHAKAR, S.C. *Economic performance of US class 1 railroads: a stochastic frontier approach*. **Applied Economics**, v.21, p.1433-1446, 1989.
- _____. *Efficiency measurement with multiple outputs and multiple inputs*. **The Journal of Productivity Analysis**, v.7, p.225-255, 1996.
- KUMBHAKAR, S.C., HJALMARSSON, L. *Labour-use efficiency in Swedish social insurance offices*. **Journal of Applied Econometrics**, v.10, p.33-47, 1995.
- LEE, L., TYLER, W.G. *The stochastic frontier production function and average efficiency: an empirical analysis*. **Journal of Econometrics**, v.7, p.385-389, 1978.
- LOVELL, C.A.K. **The measurement of productive efficiency: techniques and applications**. USA: Oxford University Press, 1993. 426 p. Cap. II, p.3-67: *Production frontiers and productive efficiency*.
- _____. *Econometric efficiency analysis: a policy-oriented review*. **European Journal of Operational Research**, v.80, p.452-461, 1995.
- _____. *Summary of some of the discussion at the Advanced Research Workshop on Efficiency Measurement*. **The Journal of Productivity Analysis**, n.7, p.341-345, 1996.

- LOVELL, C.A.K., RICHARDSON, S., TRAVERS, P. et al. *Resources and functionings: a new view of inequality in Australia*. USA: **University of North Carolina, Department of Economics, Working Paper**, n.90-8, 47 p.,1990.
- MEEUSEN, W., VAN DEN BROECK, J. *Efficiency estimation from Cobb-Douglas production functions with composed error*. **International Economic Review**, v.18, n.2, 1977.
- MENSAH, Y.M. *A simplification of the Kopp-Diewert method of decomposing cost efficiency and some implications*. **Journal of Econometrics**, v.60, 133-144, 1994.
- NICOLAU, J.A. **Eficiência técnica na produção: uma análise da agropecuária catarinense**. Florianópolis: UFSC, 1983. 75 p. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Santa Catarina, 1983.
- OLSON, J.A. SCHMIDT, P. WALDMAN, D.M. *A Monte Carlo study of estimators of stochastic frontier production functions*. **Journal of Econometrics**, v.13, p.67-82, 1980.
- PINHEIRO, A.C. *Technological progress and diffusion in Brazilian manufacturing*. **Revista Brasileira de Economia**, v.49, n.1, p.3-20, 1995.
- RAO, P. *A note on econometrics of joint production*. **Econometrica**, v.37, n.4, p.737-738, 1969.
- RICHMOND, J. *Estimating the efficiency of production*. **International Economic Review**, v.15, n.2, p.515-521, 1974.
- RUGGIERO, J. *Efficiency of educational production: an analysis of New York school districts*. **The Review of Economics and Statistics**, v.LXXVIII, n.3, p.499-509, 1996.
- _____. *A new approach for technical efficiency estimation in multiple output production*. **European Journal of Operational Research**, v.111, p.369-380,1998.
- _____. **[Standard errors of canonical coefficient]** USA: University of Dayton, 10 jun. 1999. Informação por Correio Eletrônico. Rede EPS. User name ruggiero@udayton.edu.
- RUPAREL, V.J. **The role of the underground economy in determining the transformation of command economies: allocative efficiency versus technological efficiency**. USA: Fordham University, 1995. 89 p. Tese (Doutorado em Economia) - Department of Economics, Fordham University, 1998.

- SCHMIDT, P. *On the statistical estimation of parametric frontier production functions*. **Review of Economics and Statistics**, v.58, n.2 , p.238-239, 1976.
- SCHMIDT, P., LIN, T. *Simple tests of alternative specifications in stochastic frontier models*. **Journal of Econometrics**, v.24, p.349-361, 1984.
- SCHMIDT, P., LOVELL, C.A.K. *Estimating technical and allocative inefficiency relative to stochastic production and cost frontiers*. **Journal of Econometrics**, v.9, p.343-366, 1979.
- SILVEIRA, J.S.T., LANZER, E.A., PEREIRA, M.F. *Análise econométrica de eficiência produtiva: aplicação à clínica cirúrgica do hospital da UFSC. III Workshop sobre Indicadores da Qualidade e Produtividade no Serviço Público*, Brasília,-DF, 1995.
- SILVEIRA, J.S.T., LANZER, E.A., PEREIRA, M.F. *Hospital productivity measurement as management tool of productive resources without loss of quality*. In: THE FOURTH CONFERENCE OF APORS, WITHIN IFORS, 1997, Austrália. **Anais ... Austrália: APORS'97 web site**, 1997. TD5.4.
- SPRENT, P. **Applied nonparametric statistical methods**. UK: Chapman and Hall, 1990., 259 p.
- TIMMER, C.P. *Using a probabilistic frontier production function to measure technical efficiency*. **Journal of Political Economy**, v.79, p.776-794, 1971.
- TOFALLIS, C. *Modelling best-practice frontiers when there are multiple outputs*. **The Business School, University of Hertfordshire**, UK, 1996, 1 v., p.1-7, mimeografado.
- _____. *Modelling best practice for multiple criteria*. **The Business School, University of Hertfordshire**, UK, 1998, 1 v., p.1-9, mimeografado.
- VINOD, H.D. *Econometrics of joint production*. **Econometrica**, v.36, n.2, 1968.
- _____. *Canonical ridge and econometrics of joint production*. **Journal of Econometrics**, v. 4, p.147-166, 1976.
- VITALIANO, D.F., TOREN, M. *Cost and efficiency in nursing homes: a stochastic frontier approach*. **Journal of Health Economics**, v.13, p.281-300, 1994.
- WEINSTEIN, M. *The sum of values from a normal and a truncated normal distribution*. **Technometrics**, v.6, p.104-105 e 469-470, 1964.

